SEQ 5

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:21 ; Search time 35.3333 Seconds
(without alignments)
37.723 Million cell updates/sec

Title: US-09-657-431-5
Perfect score: 75
Sequence: 1 RNPDGDVGGPWK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	75	100.0	12	22	AAB92091	Laminin fragment S
2	75	100.0	12	22	AAB36566	Mammalian kringle
3	70	93.3	11	18	AAW34297	Kringle 5 peptide
4	70	93.3	11	21	AAB01898	Human plasminogen
5	70	93.3	18	22	AAG67233	Antigenic peptide
6	70	93.3	23	22	AAB83293	Kringle peptide #2
7	70	93.3	23	22	AAB92096	Laminin fragment S
8	70	93.3	23	22	AAB36570	Mammalian kringle
9	70	93.3	24	22	AAB92090	Laminin fragment S
10	70	93.3	24	22	AAB92095	Laminin fragment S
11	70	93.3	24	22	AAB36565	Mammalian kringle
12	70	93.3	79	18	AAW19256	Human plasminogen
13	70	93.3	90	21	AAB01914	Human plasminogen
14	70	93.3	91	21	AAY58868	Human plasminogen
15	70	93.3	93	21	AAB01917	Human plasminogen
16	70	93.3	95	21	AAB01913	Human plasminogen
17	70	93.3	98	21	AAB01916	Human plasminogen
18	70	93.3	101	18	AAW34286	Human kringle 5 pe
19	70	93.3	101	18	AAW34288	Monkey kringle 5 p
20	70	93.3	101	21	AAB01890	Human plasminogen
21	70	93.3	101	21	AAB01892	Rhesus monkey plas
22	70	93.3	101	21	AAB01912	Human plasminogen
23	70	93.3	104	21	AAB01915	Human plasminogen
24	70	93.3	189	21	AAB01918	Human plasminogen
25	70	93.3	192	21	AAB01919	Human plasminogen
26	70	93.3	266	22	AAU32126	Novel human secret
27	70	93.3	266	22	AAU32129	Novel human secret
28	70	93.3	266	22	AAU32136	Novel human secret
29	70	93.3	271	21	AAB08407	A human angiogenes
30	70	93.3	357	20	AAY25408	Human tissue facto
31	70	93.3	371	13	AAR22502	[GARSYQ]-[Plasmino
32	70	93.3	380	13	AAR22504	[GARSYQ]-[Plasmino
33	70	93.3	437	19	AAW51457	Human plasminogen
34	70	93.3	453	20	AAW95051	Amino acid sequenc
35	70	93.3	467	13	AAR22499	[GARSYQ]-[Plasmino
36	70	93.3	476	13	AAR22503	[GARSYQ]-[Plasmino
37	70	93.3	566	20	AAY02100	A multifunctional
38	70	93.3	790	15	AAR60519	Human 'Glu' plasmi
39	70	93.3	790	22	AAB36562	Mammalian kringle
40	70	93.3	791	18	AAW34285	Human plasminogen.
41	70	93.3	791	21	AAB01887	Human plasminogen,
42	70	93.3	791	21	AAY99589	Human plasminogen.
43	70	93.3	791	21	AAY50867	Human plasminogen
44	70	93.3	791	22	AAG67223	Amino acid sequenc
45	70	93.3	807	13	AAR20013	PA mutant Plg 1-54

ALIGNMENTS

RESULT 1
AAB92091
ID AAB92091 standard; Peptide; 12 AA.
XX
AC AAB92091;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1267.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents

PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 610; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 75; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00069;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWK 12
||||||||||||
Db 1 RNPDGDVGGPWK 12

RESULT 2
AAB36566
ID AAB36566 standard; Peptide; 12 AA.
XX
AC AAB36566;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:5.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 5; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 75; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00069;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWK 12
||||||||||||
Db 1 RNPDGDVGGPWK 12

RESULT 3
AAW34297
ID AAW34297 standard; peptide; 11 AA.
XX
AC AAW34297;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Arg"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.
XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ, Gubbins EJ, Wang J;

```
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 4; Page 43; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   11 AA;

  Query Match             93.3%;  Score 70;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.003;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||||||
Db       1 RNPDGDVGGPW 11


RESULT 4
AAB01898
ID   AAB01898 standard; peptide; 11 AA.
XX
AC   AAB01898;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 peptide fragment #4.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   11
FT                   /note= "C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 4; Column 36; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   11 AA;

  Query Match             93.3%;  Score 70;  DB 21;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.003;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||||||
Db       1 RNPDGDVGGPW 11


RESULT 5
AAG67233
ID   AAG67233 standard; peptide; 18 AA.
XX
AC   AAG67233;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Antigenic peptide fragment of human angiostatin.
XX
KW   Angiostatin; plasminogen; sulfydryl donor; angiogenesis; tumour;
KW   angiogenic disease; neoplastic disease; connective tissue disorder;
KW   rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;
KW   diabetic retinopathy; corneal graft rejection; cardiovascular disease;
KW   cerebral vascular disease; diabetes; immune disorder;
KW   chronic inflammation; autoimmunity.
XX
OS   Homo sapiens.
XX
PN   WO200158921-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-US04021.
XX
PR   08-FEB-2000; 2000US-0500397.
XX
```

PA (NOUN) UNIV NORTHWESTERN.
XX
PI Soff G, Gately ST, Twardowski P;
XX
DR WPI; 2001-550019/61.
XX
PT Producing angiostatin for treating angiogenic diseases involves
PT contacting plasminogen with plasminogen activator and sulfhydryl donor
PT simultaneously, or producing plasmin which is contacted with sulfhydryl
PT donor -
XX
PS Example 7; Page 47; 101pp; English.
XX
CC The specification describes a method for generating angiostatin in
CC vitro. The method comprises contacting plasminogen with a sulfydryl
CC donor, or culturing cells capable of producing plasminogen activator
CC in conditioned culture medium (CCM) and contacting the CCM with
CC plasminogen. Angiostatin produced by method of the invention is useful
CC for treating animals with angiogenesis diseases. It is useful for
CC treating an angiogenic disease such as neoplastic diseases (e.g. tumours
CC and tumour metastasis), benign tumours (e.g. hemangiomas, acoustic
CC neuromas, etc), connective tissue disorders (e.g. rheumatoid arthritis
CC and atherosclerosis), ocular angiogenic diseases (e.g. diabetic
CC retinopathy, corneal graft rejection, etc), cardiovascular diseases,
CC cerebral vascular diseases, diabetes-associated diseases and immune
CC disorders (e.g. chronic inflammation and autoimmunity). The present
CC sequence represents an antigenic peptide derived from human angiostatin.
XX
SQ Sequence 18 AA;

Query Match 93.3%; Score 70; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 6
AAB83293
ID AAB83293 standard; peptide; 23 AA.
XX
AC AAB83293;
XX
DT 13-JUL-2001 (first entry)
XX
DE Kringle peptide #2.
XX
KW Pulmonary delivery; bioconjugation; pulmonary fluid protein; opioid;
KW systemic drug delivery; antihistamine; anti-angina; anti-hypertensive;
KW anti-arrhythmic; anti-depressant; bronchodilator; anti-inflammatory;
KW anti-thyroid deficiency; Kringle.
XX
OS Synthetic.
XX
PN WO200117568-A2.
XX
PD 15-MAR-2001.
XX
PF 07-SEP-2000; 2000WO-IB01429.
XX
PR 07-SEP-1999; 99US-0152681.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Ezrin AM, Fleser A, Robitaille M, Milner PG, Bridon DP;
XX
DR WPI; 2001-354657/37.
XX
PT Pulmonary delivery of therapeutic agents which are capable of forming
PT covalent bonds with amino, hydroxyl or thiol groups on pulmonary or
PT blood components -
XX
PS Example 13; Page 137; 184pp; English.
XX
CC The present invention describes a modified therapeutic agent comprising a
CC therapeutic agent and a reactive group which reacts with groups on
CC pulmonary or blood components to form a stable covalent bond, where the
CC therapeutic agent may be a peptide. Pulmonary drug delivery is useful as
CC it increases the drug retention-time in the lungs and reduces the risk of
CC extrapulmonary side effects. Modified therapeutic agents of this type may
CC be antihistamines, anti-angina, anti-hypertensive or anti-arrhythmic
CC agents, anti-depressants, bronchodilators, opioids or their analogues,
CC anti-inflammatory agents, or anti-thyroid deficiency agents. The present
CC sequence is a kringle peptide.
XX
SQ Sequence 23 AA;

Query Match 93.3%; Score 70; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 7
AAB92096
ID AAB92096 standard; Peptide; 23 AA.
XX
AC AAB92096;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1272.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 611; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic

CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 23 AA;

Query Match 93.3%; Score 70; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 8
AAB36570
ID AAB36570 standard; Peptide; 23 AA.
XX
AC AAB36570;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:9.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 5; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 23 AA;

Query Match 93.3%; Score 70; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 9
AAB92090
ID AAB92090 standard; Peptide; 24 AA.
XX
AC AAB92090;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1266.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 609; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.

CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 24 AA;

Query Match 93.3%; Score 70; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 10
AAB92095
ID AAB92095 standard; Peptide; 24 AA.
XX
AC AAB92095;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1271.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 611; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 24 AA;

Query Match 93.3%; Score 70; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
||||||||||||
Db 1 RNPDGDVGGPW 11

RESULT 11
AAB36565
ID AAB36565 standard; Peptide; 24 AA.
XX
AC AAB36565;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:4.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 5; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to

```
CC   support tumour growth). The peptides are useful for treating these
CC   diseases in mammalian or human patients. AAB36562 represents a mammalian
CC   kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC   claimed kringle 5 peptides from the present invention.
XX
SQ   Sequence   24 AA;

  Query Match             93.3%;  Score 70;  DB 22;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.0062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||
Db      1 RNPDGDVGGPW 11


RESULT 12
AAW19256
ID   AAW19256 standard; Peptide; 79 AA.
XX
AC   AAW19256;
XX
DT   27-FEB-1998  (first entry)
XX
DE   Human plasminogen Kringle 5 fragment.
XX
KW   Plasminogen; Kringle 5; cell proliferation inhibitor; angiogenesis;
KW   diagnosis; therapeutic.
XX
OS   Homo sapiens.
XX
PN   WO9723500-A1.
XX
PD   03-JUL-1997.
XX
PF   13-DEC-1996;   96WO-US20447.
XX
PR   12-DEC-1996;   96US-0763528.
PR   13-DEC-1995;   95US-0008519.
XX
PA   (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI   Cao Y,  Folkman MJ;
XX
DR   WPI; 1997-350965/32.
XX
PT   Plasminogen Kringle 5 peptide - which inhibits endothelial cell
PT   proliferation, useful to treat angiogenesis mediated diseases and in
PT   detection and diagnosis
XX
PS   Claim 1; Page 8; 51pp; English.
XX
CC   This sequence is an isolated fragment of the Kringle 5 peptide
CC   corresponding to amino acid 462 of the human plasminogen protein which
CC   can be used in a novel method to inhibit endothelial cell proliferation
CC   activity. The protein can be used to treat angiogenesis mediated
CC   diseases, e.g. haemangioma, solid tumours, leukaemia, metastasis,
CC   telangiectasia, psoriasis, scleroderma, pyogenic granuloma, myocardial
CC   angiogenesis, plaque neovascularisation, coronary or cerebral
CC   collaterals, arteriovenous malformations, ischemic limb angiogenesis,
CC   corneal diseases, rubeosis, neovascular glaucoma, diabetic retinopathy,
CC   retrolental fibroplasia, arthritis, diabetic neovascularisation,
CC   muscular degeneration, peptic ulcer, Helicobacter related disease,
CC   fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
CC   menstruation, placentation or cat scratch fever, and to stimulate wound
CC   healing. The protein and antibodies generated from it can be used to
CC   screen for agonists and antagonists or in detection, imaging and
CC   diagnosis.
XX
SQ   Sequence   79 AA;

  Query Match             93.3%;  Score 70;  DB 18;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 0.019;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||
Db     52 RNPDGDVGGPW 62


RESULT 13
AAB01914
ID   AAB01914 standard; Protein; 90 AA.
XX
AC   AAB01914;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 (Val454-Ala543).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   90 AA;

  Query Match             93.3%;  Score 70;  DB 21;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
```

```
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db     60 RNPDGDVGGPW 70


RESULT 14
AAY58868
ID   AAY58868 standard; Protein; 91 AA.
XX
AC   AAY58868;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Human plasminogen mature polypeptide.
XX
KW   Anti-angiogenic; angiogenesis inhibitor; cancer; tumour; therapy;
KW   plasminogen; human.
XX
OS   Homo sapiens.
XX
PN   WO200004052-A2.
XX
PD   27-JAN-2000.
XX
PF   16-JUL-1999;   99WO-GB02292.
XX
PR   16-JUL-1998;   98GB-0015505.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Smith RAG,  Bright JR,  Steward M,  Cox VF;
XX
DR   WPI; 2000-182406/16.
XX
PT   New soluble derivative of anti-angiogenic polypeptide useful for
PT   treatment of primary or secondary cancers, contains covalently attached
PT   membrane-binding elements for targeting  -
XX
PS   Example 3; Page 21; 36pp; English.
XX
CC   The present sequence is that of human plasminogen mature
CC   polypeptide.  The invention relates to new soluble derivatives (I)
CC   of anti-angiogenic polypeptides.  (I) comprise 2 or more
CC   heterologous membrane binding elements (MBEs, see AAY58855-61) with
CC   low membrane affinity that are covalently attached to a soluble
CC   anti-angiogenic polypeptide such as a non-catalytic region of human
CC   plasminogen.  The MBEs interact independently with thermodynamic
CC   additivity, with components of the vascular endothelium.  (I)
CC   provide targeted delivery of the anti-angiogenic polypeptide to
CC   cell membranes and sites of active angiogenesis, particularly the
CC   vascular endothelium, and therefore increase the local
CC   concentration and reduce the risk of adverse effects on normal
CC   processes elsewhere in the vasculature.  They are used in a claimed
CC   method of treatment of primary or secondary tumour.
XX
SQ   Sequence   91 AA;

  Query Match             93.3%;  Score 70;  DB 21;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db     56 RNPDGDVGGPW 66


RESULT 15
AAB01917
ID   AAB01917 standard; Protein; 93 AA.
XX
AC   AAB01917;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 (Val454-Phe546).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   93 AA;

  Query Match             93.3%;  Score 70;  DB 21;  Length 93;
  Best Local Similarity   100.0%;  Pred. No. 0.022;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db     60 RNPDGDVGGPW 70


Search completed: November  8, 2002, 09:34:51
Job time : 35.3333 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:35:26 ; Search time 16 Seconds
(without alignments)
36.034 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 6
Sequence: 1 RKLYDY 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	100.0	810	1	PLHU	plasmin (EC 3.4.21
2	6	100.0	812	1	PLMS	plasmin (EC 3.4.21
3	5	83.3	42	2	B42177	KRAB-domain-contai
4	5	83.3	72	2	D82937	conserved hypothet
5	5	83.3	79	2	D97807	ankyrin like prote
6	5	83.3	104	2	G81291	probable periplasm
7	5	83.3	107	2	F86048	hypothetical prote
8	5	83.3	107	2	B91202	hypothetical prote
9	5	83.3	110	2	G70021	hypothetical prote
10	5	83.3	112	1	B29654	proteinase inhibit
11	5	83.3	126	2	AE0091	probable flagellar
12	5	83.3	129	2	S22688	ipgA protein - Shi
13	5	83.3	177	2	S59643	hypothetical prote
14	5	83.3	181	1	A64477	L-fuculose-phospha
15	5	83.3	187	2	T47984	hypothetical prote
16	5	83.3	190	2	E64358	ribosomal protein
17	5	83.3	205	2	B82601	transcription regu
18	5	83.3	209	2	T03894	MADS box protein -
19	5	83.3	215	2	A95406	glutathione transf
20	5	83.3	231	2	T02765	glutathione transf
21	5	83.3	232	2	AI3304	response regulator
22	5	83.3	234	2	A97652	response regulator
23	5	83.3	234	2	AH2875	two component resp
24	5	83.3	234	2	A84515	probable MADS-box
25	5	83.3	234	2	T52100	MADS-box transcrip
26	5	83.3	247	2	JU0393	karasurin - Mongol
27	5	83.3	247	2	JC5032	karasurin-B - Tric
28	5	83.3	248	2	AE0281	two-component regu
29	5	83.3	264	2	A71853	hypothetical prote
30	5	83.3	264	2	G97170	aminoglycoside N3'
31	5	83.3	266	2	S59237	hypothetical prote
32	5	83.3	280	2	D71160	probable chemotaxi
33	5	83.3	283	2	G75002	chemotaxis protein
34	5	83.3	289	1	RLTZT	rRNA N-glycosidase
35	5	83.3	289	2	JC5606	karasurin C - Tric
36	5	83.3	305	2	E83852	D-alanyl-D-alanine
37	5	83.3	307	2	A71602	rifin PFB0955w - m
38	5	83.3	322	2	A70661	hypothetical prote
39	5	83.3	326	2	D90241	d-3-phosphoglycera
40	5	83.3	336	2	D86710	collagen adhesin [
41	5	83.3	368	2	F82570	DnaJ protein XF233
42	5	83.3	376	2	A90206	histidinol-phospha
43	5	83.3	376	2	E70361	chaperone DnaJ - A
44	5	83.3	379	2	S55900	DNAJ-like protein
45	5	83.3	379	2	T41633	psi protein - fiss

ALIGNMENTS

RESULT 1
PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N;Alternate names: plasminogen precursor [misnomer]
N;Contains: angiostatin; microplasmin; plasminogen
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C;Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the f
A;Reference number: A35229; MUID:90202879
A;Accession: A35229
A;Molecule type: DNA
A;Residues: 1-810 <PET>
A;Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A;Experimental source: leukocyte; lung fibroblast
R;Malgaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A;Title: Definition of the transcription initiation site of human plasminogen gene in
A;Reference number: I52242; MUID:91097523
A;Accession: I52242
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <MAL1>
A;Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613
R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human
A;Reference number: A26646; MUID:87162490
A;Accession: A26646
A;Molecule type: mRNA
A;Residues: 1-471,'D',473-810 <FOR>
A;Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
A;Experimental source: liver
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human a
A;Reference number: I45961; MUID:85023311
A;Accession: I62738
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 292-471,'D',473-810 <MAL2>
A;Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A;Accession: I84609
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 367-419 <MAL3>
A;Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg,
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of huma

A;Reference number: S03735; MUID:81212097
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71,'E',73-76 <BRU>
R;Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: protein
A;Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R;Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A;Title: Primary structure of the B-chain of human plasmin.
A;Reference number: A04627; MUID:77225245
A;Accession: A04627
A;Molecule type: protein
A;Residues: 581-810 <WI1>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla

A;Reference number: A04625; MUID:75093329
A;Accession: A04625
A;Molecule type: protein
A;Residues: 20-50,'Q',51-71,'E',73-85,87-100 <WI2>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A;Reference number: A04626; MUID:76043692
A;Accession: A04626
A;Molecule type: protein
A;Residues: 483-507,'E',509-604 <WI3>
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A;Title: The primary structure of human plasminogen. II. The histidine loop of human pla
A;Reference number: A92125; MUID:73149248
A;Contents: annotation; active site
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A;Reference number: A92048; MUID:69234739
A;Contents: annotation; active site
R;Trexler, M.; Vali, Z.; Patthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A;Reference number: A92382; MUID:82213905
A;Contents: annotation; omega-aminocarboxylic acid binding sites
R;Vali, Z.; Patthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A;Reference number: A92458; MUID:85054794
A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A;Title: Kringle domains of human angiostatin. Characterization of the anti-proliferativ
A;Reference number: A58811; MUID:97067211
A;Contents: annotation
R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A;Title: Generation of an angiostatin-like fragment from plasminogen by stromelysin-1 (M
A;Reference number: A58812; MUID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR
A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PMK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.
A;Reference number: A58819; MUID:92031502
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plas
A;Reference number: A58818; MUID:92031503
A;Contents: annotation
R;de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook,
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at
A;Reference number: A39483; MUID:92118803
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmi
A;Reference number: A58817; MUID:94237158
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH
d PIR:FGHUGB).
C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately
rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibit
C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial c
C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiostatin.
ting solid tumors.
C;Genetics:
A;Gene: GDB:PLG
A;Cross-references: GDB:119498; OMIM:173350
A;Map position: 6q26-6q27
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1;
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;20-96/Domain: activation peptide #status experimental <APT>

```
F;79-466/Product: angiostatin #status experimental <AST>
F;97-580,581-810/Product: plasmin #status experimental <MAT>
F;97-580/Domain: plasmin chain A #status experimental <CHA>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;550-580,581-810/Product: microplasmin #status experimental <MMT>

  Query Match             100.0%;  Score 6;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db     549 RKLYDY 554


RESULT 2
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiostatin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of t
A;Reference number: A38514; MUID:91184812
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DEG>
A;Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LI2>
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many ot
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, both plasmin
mediately after dissociation from the clot. In the presence of the inhibitor, the activa
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiostatin. To
eful in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiostatin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
 bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

  Query Match             100.0%;  Score 6;  DB 1;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db     549 RKLYDY 554


RESULT 3
B42177
KRAB-domain-containing zinc finger protein D19S19 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C;Accession: B42177
R;Constantinou-Deltas, C.D.; Gilbert, J.; Bartlett, R.J.; Herbstreith, M.; Roses, A.D
Genomics 12, 581-589, 1992
A;Title: The identification and characterization of KRAB-domain-containing zinc finge
A;Reference number: A42177; MUID:92217982
A;Accession: B42177
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-42 <CON>
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:95677, NCBIP:95689)
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

  Query Match             83.3%;  Score 5;  DB 2;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      25 RKLYD 29


RESULT 4
D82937
conserved hypothetical UU080 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82937
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A;Reference number: A82870
A;Accession: D82937
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-72 <GLA>
A;Cross-references: GB:AE002107; GB:AF222894; NID:g6899022; PIDN:AAF30485.1; GSPDB:GN
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: UU080
A;Genetic code: SGC3

  Query Match             83.3%;  Score 5;  DB 2;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      50 RKLYD 54


RESULT 5
D97807
ankyrin like protein [imported] - Rickettsia conorii (strain Malish 7)
```

C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97807
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: D97807
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL03398.1; PID:g15619964; GSPDB:GN00173
C;Genetics:
A;Gene: RC0860

  Query Match 83.3%; Score 5; DB 2; Length 79;
  Best Local Similarity 100.0%; Pred. No. 28;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYD 5
  |||||
Db 21 RKLYD 25

RESULT 6
G81291
probable periplasmic protein Cj1456c [imported] - Campylobacter jejuni (strain NCTC 1116
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 20-Apr-2000
C;Accession: G81291
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912
A;Accession: G81291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-104 <PAR>
A;Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73879.1; PID:g696888
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1456c
C;Superfamily: Campylobacter jejuni probable periplasmic protein Cj1456c

  Query Match 83.3%; Score 5; DB 2; Length 104;
  Best Local Similarity 100.0%; Pred. No. 35;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLYDY 6
  |||||
Db 53 KLYDY 57

RESULT 7
F86048
hypothetical protein Z5138 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F86048
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F86048
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <STO>
A;Cross-references: GB:AE005174; NID:g12518480; PIDN:AAG58850.1; GSPDB:GN00145; UWGP:Z51
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5138

  Query Match 83.3%; Score 5; DB 2; Length 107;
  Best Local Similarity 100.0%; Pred. No. 36;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYD 5
  |||||
Db 68 RKLYD 72

RESULT 8
B91202
hypothetical protein ECs4586 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: B91202
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB38009.1; PID:g13364061; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4586

  Query Match 83.3%; Score 5; DB 2; Length 107;
  Best Local Similarity 100.0%; Pred. No. 36;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYD 5
  |||||
Db 68 RKLYD 72

RESULT 9
G70021
hypothetical protein yusN - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 28-Jul-2000
C;Accession: G70021
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: G70021
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <KUN>
A;Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15275.1; PID:e11843
A;Experimental source: strain 168
C;Genetics:
A;Gene: yusN
C;Superfamily: Bacillus subtilis hypothetical protein yusN

  Query Match 83.3%; Score 5; DB 2; Length 110;
  Best Local Similarity 100.0%; Pred. No. 37;

```
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     69 RKLYD 73


RESULT 10
B29654
proteinase inhibitor (PSTI type), submandibular - lion
C;Species: Panthera leo (lion)
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C;Accession: B29654
R;Reisinger, P.W.M.; Hochstrasser, K.; Goettlicher, I.; Eulitz, M.; Wachter, E.
Biol. Chem. Hoppe-Seyler 368, 717-726, 1987
A;Title: The amino-acid sequences of the double-headed proteinase inhibitors from cat, l
A;Reference number: A94650; MUID:87299011
A;Accession: B29654
A;Molecule type: protein
A;Residues: 1-112 <REI>
C;Superfamily: submandibular proteinase inhibitor; Kazal proteinase inhibitor homology
C;Keywords: duplication; saliva; serine proteinase inhibitor; submandibular gland
F;6-60/Domain: Kazal proteinase inhibitor homology <KP1>
F;62-111/Domain: Kazal proteinase inhibitor homology <KP2>
F;8-42,20-39,28-60,64-93,71-90,79-111/Disulfide bonds: #status predicted
F;22/Inhibitory site: Lys (trypsin) #status predicted
F;73/Inhibitory site: Met (chymotrypsin, elastase) #status predicted

  Query Match             83.3%;  Score 5;  DB 1;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     53 RKLYD 57


RESULT 11
AE0091
probable flagellar protein fliS [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C;Accession: AE0091
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-126 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89592.1; PID:g15978824; GSPDB:GN00175
C;Genetics:
A;Gene: fliS
C;Superfamily: flagellar protein fliS

  Query Match             83.3%;  Score 5;  DB 2;  Length 126;
  Best Local Similarity   100.0%;  Pred. No. 41;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLYDY 6
          |||||
Db     86 KLYDY 90


RESULT 12
S22688
ipgA protein - Shigella flexneri virulence plasmid pWR100
C;Species: Shigella flexneri
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 08-Oct-1999
C;Accession: S22688
R;Allaoui, A.; Mounier, J.; Prevost, M.C.; Sansonetti, P.J.; Parsot, C.
Mol. Microbiol. 6, 1605-1616, 1992
A;Title: icsB: a Shigella flexneri virulence gene necessary for the lysis of protrusi
A;Reference number: S22687; MUID:92356824
A;Accession: S22688
A;Molecule type: DNA
A;Residues: 1-129 <ALL>
A;Cross-references: EMBL:M86530; NID:g155321; PIDN:AAD15222.1; PID:g155323
C;Genetics:
A;Gene: ipgA
A;Genome: plasmid

  Query Match             83.3%;  Score 5;  DB 2;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db      3 RKLYD 7


RESULT 13
S59643
hypothetical protein SPAC13C5.06c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 10-Dec-1999
C;Accession: S59643; T37619
R;Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, July 1995
A;Reference number: S58093
A;Accession: S59643
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <DEV>
A;Cross-references: EMBL:Z50112; NID:g908889; PIDN:CAA90457.1; PID:g908895
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z21731
A;Accession: T37619
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-177 <DE2>
A;Cross-references: EMBL:Z50112; PIDN:CAA90457.1; GSPDB:GN00066; SPDB:SPAC13C5.06c
A;Experimental source: strain 972h-; cosmid c13C5
C;Genetics:
A;Gene: SPDB:SPAC13C5.06c
A;Map position: 1
A;Introns: 76/1
C;Superfamily: Schizosaccharomyces SPAC13C5.06c

  Query Match             83.3%;  Score 5;  DB 2;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLYDY 6
          |||||
Db     70 KLYDY 74


RESULT 14
A64477
L-fuculose-phosphate aldolase homolog - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: A64477
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999
A;Accession: A64477
```

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-181 <BUL>
A;Cross-references: GB:U67582; GB:L77117; NID:g1592064; PIDN:AAB99428.1; PID:g1592067; T
C;Genetics:
A;Map position: FOR1381152-1381697
C;Superfamily: L-ribulose-phosphate 4-epimerase

Query Match 83.3%; Score 5; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYD 5
|||||
Db 11 RKLYD 15

RESULT 15
T47984
hypothetical protein F21F14.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000
C;Accession: T47984
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24481
A;Accession: T47984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-187 <CHO>
A;Cross-references: EMBL:AL138642
A;Experimental source: cultivar Columbia; BAC clone F21F14
C;Genetics:
A;Map position: 3
A;Note: F21F14.90
C;Superfamily: Arabidopsis thaliana hypothetical protein F21F14.90

Query Match 83.3%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYD 5
|||||
Db 13 RKLYD 17

Search completed: November 8, 2002, 09:38:01
Job time : 17 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:34:56 ; Search time 11 Seconds
(without alignments)
21.120 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 6
Sequence: 1 RKLYDY 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	100.0	810	1	PLMN_HUMAN	P00747 homo sapien
2	6	100.0	812	1	PLMN_MOUSE	P20918 mus musculu
3	5	83.3	112	1	IPSG_PANLE	P08481 panthera le
4	5	83.3	129	1	IPGA_SHIFL	P33547 shigella fl
5	5	83.3	177	1	YA16_SCHPO	Q09688 schizosacch
6	5	83.3	181	1	YE18_METJA	Q58813 methanococc
7	5	83.3	190	1	RL5_METJA	P54040 methanococc
8	5	83.3	266	1	YGS4_YEAST	P46947 saccharomyc
9	5	83.3	289	1	RIPS_TRIKI	P24478 trichosanth
10	5	83.3	289	1	RIPT_TRIKI	P09989 trichosanth
11	5	83.3	302	1	RRM1_DROME	Q9vep1 drosophila
12	5	83.3	305	1	DDL_BACHD	Q9kcf0 bacillus ha
13	5	83.3	368	1	DNAJ_XYLFA	Q9pb06 xylella fas
14	5	83.3	376	1	DNJ2_AQUAE	O66921 aquifex aeo
15	5	83.3	376	1	HIS8_SULSO	O33770 sulfolobus
16	5	83.3	379	1	PSI_SCHPO	Q09912 schizosacch
17	5	83.3	387	1	YRS8_CAEEL	Q10004 caenorhabdi
18	5	83.3	389	1	DNAJ_MYCGE	P47265 mycoplasma
19	5	83.3	405	1	DCDA_HELPJ	Q9zme5 helicobacte
20	5	83.3	405	1	DCDA_HELPY	P56129 helicobacte
21	5	83.3	458	1	ERR3_HUMAN	O75454 homo sapien
22	5	83.3	488	1	EXON_HSV6U	P24447 human herpe
23	5	83.3	488	1	EXON_HSV6Z	P52448 human herpe
24	5	83.3	525	1	DIMH_CAEEL	O17397 caenorhabdi
25	5	83.3	527	1	YBIP_ECOLI	P75785 escherichia
26	5	83.3	692	1	VP3_ROTPC	P26192 porcine rot
27	5	83.3	974	1	CC15_YEAST	P27636 saccharomyc
28	5	83.3	1161	1	RPO2_FOWPV	Q9j544 fowlpox vir
29	5	83.3	1174	1	ZO2_CANFA	Q95168 canis famil
30	5	83.3	1193	1	ACE_CHICK	Q10751 gallus gall
31	5	83.3	1337	1	P152_YEAST	P39685 saccharomyc
32	5	83.3	2022	1	ANT1_ONCVO	P21249 onchocerca
33	5	83.3	3951	1	VGF1_IBVB	P27920 avian infec
34	4	66.7	45	1	RIP2_TRIKI	P23029 trichosanth
35	4	66.7	45	1	Y00D_BPT4	Q01434 bacteriopha
36	4	66.7	46	1	LHA1_ECTHA	P80100 ectothiorho
37	4	66.7	52	1	NAHQ_PSEPU	Q51946 pseudomonas
38	4	66.7	62	1	TS24_DENJA	P01407 dendroaspis
39	4	66.7	62	1	TS81_DENAN	P01410 dendroaspis
40	4	66.7	62	1	TS82_DENAN	P01411 dendroaspis
41	4	66.7	63	1	TS91_DENAN	P01408 dendroaspis
42	4	66.7	63	1	TS92_DENAN	P01409 dendroaspis
43	4	66.7	67	1	YMOA_YERPE	P27720 yersinia pe
44	4	66.7	71	1	YAAA_BACSU	P05650 bacillus su
45	4	66.7	71	1	YDGT_ECOLI	P76179 escherichia

ALIGNMENTS

RESULT 1
PLMN_HUMAN
ID PLMN_HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;

RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vali Z., Patthy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vali Z., Patthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RT are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RT plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE 268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevich R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;
RL Submitted (JUN-1995) to the PDB data bank.
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.
RX MEDLINE=91095410; PubMed=1986355;
RA Ichinose A., Espling E.S., Takamatsu J., Saito H., Shinmyozu K.,

```
  Query Match             100.0%;  Score 6;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 2
PLMN_MOUSE
ID   PLMN_MOUSE     STANDARD;      PRT;   812 AA.
AC   P20918;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN   PLG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91184812; PubMed=2081600;
RA   Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse plasminogen and
RT   localization of the gene to mouse chromosome 17.";
RL   Genomics 8:49-61(1990).
RN   [2]
RP   CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX   MEDLINE=95042728; PubMed=7525077;
RA   O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA   Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT   "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT   suppression of metastases by a Lewis lung carcinoma.";
RL   Cell 79:315-328(1994).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC       NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC       METASTATIC TUMORS IN VIVO.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04766; AAA50168.1; -.
DR   PIR; A38514; A38514.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   MGD; MGI:97620; Plg.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
FT   CHAIN        20    581       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE      20     97       ACTIVATION PEPTIDE.
FT   CHAIN        98    581       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN        98   ?436       ANGIOSTATIN.
FT   CHAIN       582    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN      103    181       KRINGLE 1.
FT   DOMAIN      184    262       KRINGLE 2.
FT   DOMAIN      275    352       KRINGLE 3.
FT   DOMAIN      377    454       KRINGLE 4.
FT   DOMAIN      481    560       KRINGLE 5.
FT   DOMAIN      582    812       SERINE PROTEASE.
FT   ACT_SITE    624    624       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    667    667       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    762    762       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     49     73       BY SIMILARITY.
FT   DISULFID     53     61       BY SIMILARITY.
FT   DISULFID    103    181       BY SIMILARITY.
FT   DISULFID    124    164       BY SIMILARITY.
FT   DISULFID    152    176       BY SIMILARITY.
FT   DISULFID    185    262       BY SIMILARITY.
FT   DISULFID    188    316       BY SIMILARITY.
FT   DISULFID    206    245       BY SIMILARITY.
FT   DISULFID    234    257       BY SIMILARITY.
FT   DISULFID    275    352       BY SIMILARITY.
FT   DISULFID    296    335       BY SIMILARITY.
FT   DISULFID    324    347       BY SIMILARITY.
FT   DISULFID    377    454       BY SIMILARITY.
FT   DISULFID    398    437       BY SIMILARITY.
FT   DISULFID    426    449       BY SIMILARITY.
FT   DISULFID    481    560       BY SIMILARITY.
FT   DISULFID    502    543       BY SIMILARITY.
FT   DISULFID    531    555       BY SIMILARITY.
FT   DISULFID    568    687       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    578    586       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    609    625       BY SIMILARITY.
FT   DISULFID    701    768       BY SIMILARITY.
FT   DISULFID    731    747       BY SIMILARITY.
FT   DISULFID    758    786       BY SIMILARITY.
SQ   SEQUENCE   812 AA;  90846 MW;  D34A74A4FC2256F8 CRC64;

  Query Match             100.0%;  Score 6;  DB 1;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554
```

```
RESULT 3
IPSG_PANLE
ID   IPSG_PANLE     STANDARD;      PRT;   112 AA.
AC   P08481;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Double-headed protease inhibitor, submandibular gland.
OS   Panthera leo (Lion).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Panthera.
OX   NCBI_TaxID=9689;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Submandibular gland;
RX   MEDLINE=87299011; PubMed=3304339;
RA   Reisinger P.W.M., Hochstrasser K., Gottlicher I., Eulitz M.,
RA   Wachter E.;
RT   "The amino-acid sequences of the double-headed proteinase inhibitors
RT   from cat, lion and dog submandibular glands.";
RL   Biol. Chem. Hoppe-Seyler 368:717-726(1987).
CC   -!- FUNCTION: THIS INHIBITOR IS COMPOSED OF TWO HOMOLOGOUS ACTIVELY
CC       INHIBITING HALVES: ONE WHICH INHIBITS TRYPSIN, THE OTHER WHICH
CC       INHIBITS ELASTASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: CONTAINS 2 KAZAL-LIKE DOMAINS.
DR   PIR; B29654; B29654.
DR   HSSP; P05586; 4OVO.
DR   InterPro; IPR001239; Kazal_inhib.
DR   InterPro; IPR002350; kazal.
DR   Pfam; PF00050; kazal; 2.
DR   PRINTS; PR00290; KAZALINHBTR.
DR   SMART; SM00280; KAZAL; 2.
DR   PROSITE; PS00282; KAZAL; 2.
KW   Serine protease inhibitor; Repeat; Submandibular gland.
FT   DOMAIN        1     60       KAZAL-LIKE 1.
FT   DOMAIN       61    112       KAZAL-LIKE 2.
FT   ACT_SITE     22     23       REACTIVE BOND 1 (TRYPSIN).
FT   ACT_SITE     73     74       REACTIVE BOND 2 (ELASTASE).
FT   DISULFID      8     42       BY SIMILARITY.
FT   DISULFID     20     39       BY SIMILARITY.
FT   DISULFID     28     60       BY SIMILARITY.
FT   DISULFID     64     93       BY SIMILARITY.
FT   DISULFID     71     90       BY SIMILARITY.
FT   DISULFID     79    111       BY SIMILARITY.
SQ   SEQUENCE   112 AA;  12740 MW;  D8062796D3FC611C CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     53 RKLYD 57


RESULT 4
IPGA_SHIFL
ID   IPGA_SHIFL     STANDARD;      PRT;   129 AA.
AC   P33547;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   IpgA protein.
GN   IPGA.
OS   Shigella flexneri.
OG   Plasmid 210 kb invasion pWR100.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M90T / SEROTYPE 5;
RX   MEDLINE=92356824; PubMed=1495389;
RA   Allaoui A., Mounier J., Prevost M.-C., Sansonetti P.J., Parsot C.;
RT   "icsB: a Shigella flexneri virulence gene necessary for the lysis of
RT   protrusions during intercellular spread.";
RL   Mol. Microbiol. 6:1605-1616(1992).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M86530; AAD15222.1; -.
DR   PIR; S22688; S22688.
KW   Plasmid; Virulence.
SQ   SEQUENCE   129 AA;  15085 MW;  8780F5B428A78847 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db      3 RKLYD 7


RESULT 5
YA16_SCHPO
ID   YA16_SCHPO     STANDARD;      PRT;   177 AA.
AC   Q09688;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Hypothetical 20.1 kDa protein C13C5.06C in chromosome I precursor.
GN   SPAC13C5.06C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z50112; CAA90457.1; -.
KW   Hypothetical protein; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    177       HYPOTHETICAL PROTEIN C13C5.06C.
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   177 AA;  20058 MW;  612DE3D3F7B64027 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLYDY 6
          |||||
Db     70 KLYDY 74


RESULT 6
```

```
YE18_METJA
ID   YE18_METJA     STANDARD;      PRT;   181 AA.
AC   Q58813;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical aldolase class II protein MJ1418.
GN   MJ1418.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ARAD/FUCA
CC       SUBFAMILY. STRONG, TO E.COLI YGBL AND H.INFLUENZAE HI1012.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U67582; AAB99428.1; -.
DR   HSSP; P11550; 3FUA.
DR   TIGR; MJ1418; -.
DR   InterPro; IPR001303; Aldolase_II.
DR   Pfam; PF00596; Aldolase_II; 1.
KW   Hypothetical protein; Zinc; Complete proteome.
FT   METAL        68     68       ZINC (BY SIMILARITY).
FT   METAL        87     87       ZINC (BY SIMILARITY).
FT   METAL        89     89       ZINC (BY SIMILARITY).
FT   METAL       147    147       ZINC (BY SIMILARITY).
SQ   SEQUENCE   181 AA;  20470 MW;  E5F3BF13722145B0 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 181;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     11 RKLYD 15


RESULT 7
RL5_METJA
ID   RL5_METJA      STANDARD;      PRT;   190 AA.
AC   P54040;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L5P.
GN   RPL5P OR MJ0469.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U67497; AAB98458.1; -.
DR   TIGR; MJ0469; -.
DR   InterPro; IPR002132; Ribosomal_L5.
DR   Pfam; PF00281; Ribosomal_L5; 1.
DR   Pfam; PF00673; Ribosomal_L5_C; 1.
DR   ProDom; PD001076; Ribosomal_L5; 1.
DR   PROSITE; PS00358; RIBOSOMAL_L5; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   190 AA;  22169 MW;  BD05366973C3A533 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 190;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLYDY 6
          |||||
Db     96 KLYDY 100


RESULT 8
YGS4_YEAST
ID   YGS4_YEAST     STANDARD;      PRT;   266 AA.
AC   P46947;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Hypothetical 30.5 kDa protein in SAE2-KEM1 intergenic region.
GN   YGL174W OR G1642.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE=96109931; PubMed=8619317;
RA   Bertani I., Coglievina M., Zaccaria P., Klima R., Bruschi C.V.;
RT   "The sequence of an 11.1 kb fragment on the left arm of Saccharomyces
RT   cerevisiae chromosome VII reveals six open reading frames including
RT   NSP49, KEM1 and four putative new genes.";
RL   Yeast 11:1187-1194(1995).
CC   -!- SIMILARITY: SOME, TO C.ELEGANS R08D7.1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X84705; CAA59179.1; -.
DR   EMBL; Z72696; CAA96886.1; -.
DR   SGD; S0003142; YGL174W.
KW   Hypothetical protein.
SQ   SEQUENCE   266 AA;  30472 MW;  1584B390121325CD CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 266;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db     212 RKLYD 216


RESULT 9
RIPS_TRIKI
ID   RIPS_TRIKI     STANDARD;      PRT;   289 AA.
AC   P24478;
DT   01-MAR-1992 (Rel. 21, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ribosome-inactivating protein karasurin precursor (rRNA
DE   N-glycosidase) (EC 3.2.2.22).
OS   Trichosanthes kirilowii (Mongolian snake-gourd).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX   NCBI_TaxID=3677;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Root tuber;
RX   MEDLINE=97356562; PubMed=9212998;
RA   Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT   "Cloning and bacterial expression of a gene encoding ribosome-
RT   inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT   kirilowii var. japonica.";
RL   Biol. Pharm. Bull. 20:711-713(1997).
RN   [2]
RP   SEQUENCE OF 24-270.
RX   MEDLINE=92005921; PubMed=1914000;
RA   Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT   "The complete amino acid sequence of an abortifacient protein,
RT   karasurin.";
RL   Chem. Pharm. Bull. 39:1244-1249(1991).
CC   -!- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
CC       60S RIBOSOMAL SUBUNITS.
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC       specific adenosine on the 28S rRNA.
CC   -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
CC       RIP SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB000666; BAA21786.1; -.
DR   PIR; JU0393; JU0393.
DR   HSSP; P09989; 1MRJ.
DR   InterPro; IPR001574; RIP.
DR   Pfam; PF00161; RIP; 1.
DR   PRINTS; PR00396; SHIGARICIN.
DR   PROSITE; PS00275; SHIGA_RICIN; 1.
KW   Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    270       KARASURIN-C.
FT   CHAIN        24    270       KARASURIN-A.
FT   PROPEP      271    289       REMOVED IN MATURE FORM.
FT   ACT_SITE    183    183       BY SIMILARITY.
SQ   SEQUENCE   289 AA;  31704 MW;  883D3E3242887B26 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      52 RKLYD 56


RESULT 10
RIPT_TRIKI
ID   RIPT_TRIKI     STANDARD;      PRT;   289 AA.
AC   P09989;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ribosome-inactivating protein alpha-trichosanthin precursor
DE   (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).
OS   Trichosanthes kirilowii (Mongolian snake-gourd).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX   NCBI_TaxID=3677;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAXIMOWICZ;
RX   MEDLINE=91153657; PubMed=1999291;
RA   Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT   "Cloning of trichosanthin cDNA and its expression in Escherichia
RT   coli.";
RL   Gene 97:267-272(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAXIMOWICZ; TISSUE=Leaf;
RX   MEDLINE=90256790; PubMed=2341400;
RA   Chow T., Feldman R.A., Lovett M., Piatak M.;
RT   "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
RT   type I ribosome-inactivating protein.";
RL   J. Biol. Chem. 265:8670-8674(1990).
RN   [3]
RP   SEQUENCE OF 24-270.
RC   STRAIN=MAXIMOWICZ; TISSUE=Tuberous root;
RX   MEDLINE=90256789; PubMed=2341399;
RA   Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,
RA   Wu P., Hwang K., Piatak M.;
RT   "Primary amino acid sequence of alpha-trichosanthin and molecular
RT   models for abrin A-chain and alpha-trichosanthin.";
RL   J. Biol. Chem. 265:8665-8669(1990).
RN   [4]
RP   SEQUENCE OF 24-270.
RC   TISSUE=Tuberous root;
RA   Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
RA   Tian G.Y., Ni C.Z.;
RT   "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and
RT   application.";
RL   Pure Appl. Chem. 58:789-798(1986).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RX   MEDLINE=94344957; PubMed=8066085;
RA   Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT   "Structure of trichosanthin at 1.88-A resolution.";
RL   Proteins 19:4-13(1994).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX   MEDLINE=95344383; PubMed=7619070;
RA   Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT   "Studies on crystal structures, active-centre geometry and
RT   depurinating mechanism of two ribosome-inactivating proteins.";
```

```
RL   Biochem. J. 309:285-298(1995).
CC   -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC       CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
CC       INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC       specific adenosine on the 28S rRNA.
CC   -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
CC       RIP SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M34858; AAA34207.1; -.
DR   EMBL; J05434; AAA34206.1; -.
DR   PIR; JT0003; RLTZT.
DR   PIR; JT0566; JT0566.
DR   PIR; A36273; A36273.
DR   PIR; A36274; A36274.
DR   PDB; 1MRJ; 07-FEB-95.
DR   PDB; 1MRK; 07-FEB-95.
DR   PDB; 1TCS; 10-JUL-95.
DR   InterPro; IPR001574; RIP.
DR   Pfam; PF00161; RIP; 1.
DR   PRINTS; PR00396; SHIGARICIN.
DR   PROSITE; PS00275; SHIGA_RICIN; 1.
KW   Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
KW   3D-structure.
FT   SIGNAL        1     23
FT   CHAIN        24    270       RIBOSOME-INACTIVATING PROTEIN ALPHA-
FT                                 TRICHOSANTHIN.
FT   PROPEP      271    289       MISSING IN MATURE PROTEIN.
FT   ACT_SITE    183    183       BY SIMILARITY.
FT   CONFLICT     57     60       IPLL -> LPLI (IN REF. 4).
FT   CONFLICT     82     84       MISSING (IN REF. 4).
FT   CONFLICT     87     87       I -> L (IN REF. 4).
FT   CONFLICT     92     92       V -> VDAGLPRNAVL (IN REF. 4).
FT   CONFLICT    143    144       KI -> GL (IN REF. 4).
FT   CONFLICT    196    196       K -> S (IN REF. 4).
FT   CONFLICT    214    216       SWS -> SLWL (IN REF. 4).
FT   CONFLICT    231    231       Q -> T (IN REF. 4).
FT   CONFLICT    234    234       S -> T (IN REF. 2).
FT   CONFLICT    246    266       MISSING (IN REF. 4).
FT   CONFLICT    247    247       T -> M (IN REF. 2).
SQ   SEQUENCE   289 AA;  31676 MW;  5CE09BB630575BB9 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      52 RKLYD 56


RESULT 11
RRM1_DROME
ID   RRM1_DROME     STANDARD;      PRT;   302 AA.
AC   Q9VEP1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ribosomal RNA methyltransferase CG5220 (EC 2.1.1.-) (rRNA
DE   (uridine-2'-O-)-methyltransferase).
GN   CG5220.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + RRNA = S-ADENOSYL-L-
CC       HOMOCYSTEINE + RRNA CONTAINING 2'-O-METHYLURIDINE.
CC   -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RRMJ
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003716; AAF55380.1; -.
DR   FlyBase; FBgn0038471; CG5220.
DR   InterPro; IPR002877; FtsJ.
DR   Pfam; PF01728; FtsJ; 1.
KW   Hypothetical protein; rRNA processing; Transferase; Methyltransferase.
SQ   SEQUENCE   302 AA;  33402 MW;  EDFB3BAD445693DA CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 302;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      61 RKLYD 65


RESULT 12
```

```
DDL_BACHD
ID   DDL_BACHD      STANDARD;      PRT;   305 AA.
AC   Q9KCF0;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylalanine synthetase)
DE   (D-Ala-D-Ala ligase).
GN   DDL OR DDLA OR BH1621.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- FUNCTION: Cell wall formation (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + D-alanine + D-alanine = ADP + phosphate
CC       + D-alanyl-D-alanine.
CC   -!- PATHWAY: ALONG WITH ALANINE RACEMASE, IT MAKES UP THE D-ALANINE
CC       BRANCH OF THE PEPTIDOGLYCAN BIOSYNTHETIC ROUTE.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE D-ALANINE--D-ALANINE LIGASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001512; BAB05340.1; -.
DR   HSSP; P07862; 1IOW.
DR   InterPro; IPR000291; Dala_Dala_ligase.
DR   Pfam; PF01820; Dala_Dala_ligas; 2.
DR   PROSITE; PS00843; DALA_DALA_LIGASE_1; 1.
DR   PROSITE; PS00844; DALA_DALA_LIGASE_2; FALSE_NEG.
KW   Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
SQ   SEQUENCE   305 AA;  33125 MW;  C806D03993D3D081 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 305;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 KLYDY 6
           |||||
Db     204 KLYDY 208


RESULT 13
DNAJ_XYLFA
ID   DNAJ_XYLFA      STANDARD;      PRT;   368 AA.
AC   Q9PB06;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chaperone protein dnaJ.
GN   DNAJ OR XF2339.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
CC   -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC       THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC   -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 CR DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004044; AAF85138.1; -.
DR   InterPro; IPR003095; DnaJ.
DR   InterPro; IPR002939; DnaJ_C.
DR   InterPro; IPR001305; DnaJ_CXXCXGXG.
DR   InterPro; IPR001623; DnaJ_N.
DR   Pfam; PF01556; DnaJ_C; 1.
DR   Pfam; PF00684; DnaJ_CXXCXGXG; 1.
DR   PRINTS; PR00625; DNAJPROTEIN.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS00636; DNAJ_1; FALSE_NEG.
DR   PROSITE; PS50076; DNAJ_2; 1.
DR   PROSITE; PS00637; DNAJ_CXXCXGXG; 1.
KW   Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
KW   Complete proteome.
FT   DOMAIN        5     70       J-DOMAIN.
FT   DOMAIN       77    111       GLY-RICH.
FT   REPEAT      137    144       CXXCXGXG MOTIF.
FT   REPEAT      153    160       CXXCXGXG MOTIF.
FT   REPEAT      175    182       CXXCXGXG MOTIF.
FT   REPEAT      189    196       CXXCXGXG MOTIF.
FT   METAL       137    137       ZINC 1 (BY SIMILARITY).
FT   METAL       140    140       ZINC 1 (BY SIMILARITY).
FT   METAL       153    153       ZINC 2 (BY SIMILARITY).
FT   METAL       156    156       ZINC 2 (BY SIMILARITY).
FT   METAL       175    175       ZINC 2 (BY SIMILARITY).
FT   METAL       178    178       ZINC 2 (BY SIMILARITY).
FT   METAL       189    189       ZINC 1 (BY SIMILARITY).
FT   METAL       192    192       ZINC 1 (BY SIMILARITY).
```

```
SQ   SEQUENCE   368 AA;  40372 MW;  8631363433B805B9 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 368;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      63 RKLYD 67


RESULT 14
DNJ2_AQUAE
ID   DNJ2_AQUAE     STANDARD;      PRT;   376 AA.
AC   O66921;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chaperone protein dnaJ-2.
GN   DNAJ2 OR AQ_703.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC       THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC   -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 CR DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000703; AAC06881.1; -.
DR   HSSP; P08622; 1XBL.
DR   InterPro; IPR003095; DnaJ.
DR   InterPro; IPR002939; DnaJ_C.
DR   InterPro; IPR001305; DnaJ_CXXCXGXG.
DR   InterPro; IPR001623; DnaJ_N.
DR   Pfam; PF00226; DnaJ; 1.
DR   Pfam; PF01556; DnaJ_C; 1.
DR   Pfam; PF00684; DnaJ_CXXCXGXG; 1.
DR   PRINTS; PR00625; DNAJPROTEIN.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS00636; DNAJ_1; 1.
DR   PROSITE; PS50076; DNAJ_2; 1.
DR   PROSITE; PS00637; DNAJ_CXXCXGXG; FALSE_NEG.
KW   Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
KW   Complete proteome.
FT   DOMAIN        8     72       J-DOMAIN.
FT   REPEAT      156    163       CXXCXGXG MOTIF.
FT   REPEAT      172    179       CXXCXGXG MOTIF.
FT   REPEAT      194    201       CXXCXGXG MOTIF.
FT   REPEAT      207    214       CXXCXGXG MOTIF.
FT   METAL       156    156       ZINC 1 (BY SIMILARITY).
FT   METAL       159    159       ZINC 1 (BY SIMILARITY).
FT   METAL       172    172       ZINC 2 (BY SIMILARITY).
FT   METAL       175    175       ZINC 2 (BY SIMILARITY).
FT   METAL       194    194       ZINC 2 (BY SIMILARITY).
FT   METAL       197    197       ZINC 2 (BY SIMILARITY).
FT   METAL       207    207       ZINC 1 (BY SIMILARITY).
FT   METAL       210    210       ZINC 1 (BY SIMILARITY).
SQ   SEQUENCE   376 AA;  41963 MW;  9C8FBFB29A8A016D CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      65 RKLYD 69


RESULT 15
HIS8_SULSO
ID   HIS8_SULSO     STANDARD;      PRT;   376 AA.
AC   O33770;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Histidinol-phosphate aminotransferase (EC 2.6.1.9) (Imidazole acetol-
DE   phosphate transaminase).
GN   HISC OR SSO0592 OR C08_058.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=97352708; PubMed=9209067;
RA   Charlebois R.L., Sensen C.W., Doolittle W.F., Brown J.R.;
RT   "Evolutionary analysis of the hisCGABdFDEHI gene cluster from the
RT   archaeon Sulfolobus solfataricus P2.";
RL   J. Bacteriol. 179:4429-4432(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=20165948; PubMed=10701121;
RA   Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
RA   Confalonieri F., Curtis B., Duguet M., Erauso G., Faguy D.,
RA   Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
RA   Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA   St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
RA   Ragan M.A., Sensen C.W.;
RT   "Gene content and organization of a 281-kbp contig from the genome of
RT   the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
RL   Genome 43:116-136(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC   -!- CATALYTIC ACTIVITY: L-histidinol-phosphate + 2-oxoglutarate = 3-
CC       (imidazol-4-yl)-2-oxopropyl phosphate + L-glutamate.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC   -!- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC   -!- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U82227; AAB63018.1; -.
DR   EMBL; Y18930; CAB57708.1; -.
DR   EMBL; AE006689; AAK40904.1; -.
DR   InterPro; IPR001511; Aminotran_1.
DR   InterPro; IPR001917; Aminotransf_2.
DR   Pfam; PF00155; aminotran_1_2; 1.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW   Histidine biosynthesis; Transferase; Aminotransferase;
KW   Pyridoxal phosphate; Complete proteome.
FT   BINDING     240    240       PYRIDOXAL PHOSPHATE (PROBABLE).
SQ   SEQUENCE   376 AA;  43193 MW;  44CAB23CD2866405 CRC64;

  Query Match             83.3%;  Score 5;  DB 1;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db     342 RKLYD 346


Search completed: November  8, 2002, 09:37:38
Job time : 12 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:36:17 ; Search time 25 Seconds
(without alignments)
41.519 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 6
Sequence: 1 RKLYDY 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	100.0	334	6	O46507	O46507 papio hamad
2	6	100.0	810	4	Q15146	Q15146 homo sapien
3	6	100.0	812	11	Q9R0W3	Q9r0w3 rattus norv
4	6	100.0	812	11	Q91WJ5	Q91wj5 mus musculu
5	5	83.3	72	16	Q9PR65	Q9pr65 ureaplasma
6	5	83.3	73	2	Q9X3B8	Q9x3b8 pediococcus
7	5	83.3	79	16	Q92HB1	Q92hb1 rickettsia
8	5	83.3	104	16	Q9PMK4	Q9pmk4 campylobact
9	5	83.3	107	2	O52124	O52124 escherichia
10	5	83.3	110	16	O32180	O32180 bacillus su
11	5	83.3	111	5	Q9VAW7	Q9vaw7 drosophila
12	5	83.3	111	5	Q9VXI6	Q9vxi6 drosophila
13	5	83.3	129	4	Q9UNJ4	Q9unj4 homo sapien
14	5	83.3	148	16	Q98F36	Q98f36 rhizobium l
15	5	83.3	157	10	Q9AUJ1	Q9auj1 oryza sativ
16	5	83.3	158	9	Q37913	Q37913 bacteriopha
17	5	83.3	170	2	Q9KX31	Q9kx31 streptococc
18	5	83.3	187	10	Q9M273	Q9m273 arabidopsis
19	5	83.3	205	16	Q9PBQ5	Q9pbq5 xylella fas
20	5	83.3	208	10	Q9XGJ5	Q9xgj5 gnetum gnem
21	5	83.3	209	10	Q40702	Q40702 oryza sativ
22	5	83.3	209	10	Q9AR51	Q9ar51 zea mays (m
23	5	83.3	212	10	Q9AR50	Q9ar50 zea mays (m
24	5	83.3	212	10	Q9AR49	Q9ar49 zea mays (m
25	5	83.3	215	16	Q9F5P2	Q9f5p2 rhizobium m
26	5	83.3	231	10	O65032	O65032 oryza sativ
27	5	83.3	231	16	Q98HL6	Q98hl6 rhizobium l
28	5	83.3	232	2	Q9ZHS1	Q9zhs1 brucella ab
29	5	83.3	232	2	O86246	O86246 helicobacte
30	5	83.3	233	16	Q9FD95	Q9fd95 rhizobium m
31	5	83.3	234	10	Q9SI38	Q9si38 arabidopsis
32	5	83.3	234	10	O49351	O49351 arabidopsis
33	5	83.3	237	2	Q9RNH4	Q9rnh4 rhodobacter
34	5	83.3	237	2	Q49198	Q49198 mycoplasma
35	5	83.3	247	10	Q9LRE3	Q9lre3 trichosanth
36	5	83.3	256	10	Q9LVC4	Q9lvc4 arabidopsis
37	5	83.3	264	16	Q9ZK74	Q9zk74 helicobacte
38	5	83.3	264	16	Q97H17	Q97h17 clostridium
39	5	83.3	266	2	Q9KX34	Q9kx34 streptococc
40	5	83.3	269	5	O97118	O97118 drosophila
41	5	83.3	277	5	Q9W0A8	Q9w0a8 drosophila
42	5	83.3	280	17	O58194	O58194 pyrococcus
43	5	83.3	283	17	Q9UYF1	Q9uyf1 pyrococcus
44	5	83.3	289	10	Q41216	Q41216 trichosanth
45	5	83.3	289	10	Q94KE4	Q94ke4 trichosanth

ALIGNMENTS

RESULT 1
O46507
ID O46507 PRELIMINARY; PRT; 334 AA.
AC O46507;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PLASMINOGEN (FRAGMENT).
GN BABPEPSG.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF029692; AAB97887.1; -.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.

```
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  36791 MW;  C7DC06E03B965286 CRC64;

  Query Match             100.0%;  Score 6;  DB 6;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db      73 RKLYDY 78


RESULT 2
Q15146
ID   Q15146      PRELIMINARY;      PRT;   810 AA.
AC   Q15146;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA   Mitchell D., Robinson J.H.;
RT   "Expression of recombinant human plasminogen and aglycoplasminogen in
RT   HeLa cells.";
RL   Fibrinolysis 0:0-0(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; M74220; AAA36451.1; -.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    810       PLASMINOGEN.
SQ   SEQUENCE   810 AA;  90555 MW;  B05C7D4B0D020B3C CRC64;

  Query Match             100.0%;  Score 6;  DB 4;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db     549 RKLYDY 554


RESULT 3
Q9R0W3
ID   Q9R0W3      PRELIMINARY;      PRT;   812 AA.
AC   Q9R0W3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PROTEIN PRECURSOR (EC 3.4.21.7).
GN   PLASMINOGEN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Bangert K., Johnsen A.H., Thorsen S.;
RT   "Rat plasminogen: cDNA and gene structure.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=91250378; PubMed=1645711;
RA   Kanalas J.J., Makker S.P.;
RT   "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT   receptor site for plasminogen.";
RL   J. Biol. Chem. 266:10825-10829(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AJ242649; CAB46014.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
SQ   SEQUENCE   812 AA;  90535 MW;  8C703C51410EBC9E CRC64;

  Query Match             100.0%;  Score 6;  DB 11;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db     549 RKLYDY 554


RESULT 4
Q91WJ5
ID   Q91WJ5      PRELIMINARY;      PRT;   812 AA.
AC   Q91WJ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014773; AAH14773.1; -.
SQ   SEQUENCE   812 AA;  90781 MW;  24173260E6A2FFD2 CRC64;

  Query Match             100.0%;  Score 6;  DB 11;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 5
Q9PR65
ID   Q9PR65      PRELIMINARY;      PRT;    72 AA.
AC   Q9PR65;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL PROTEIN UU080.
GN   UU080.
OS   Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Ureaplasma.
OX   NCBI_TaxID=134821;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SEROVAR 3;
RX   MEDLINE=20500219; PubMed=11048724;
RA   Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA   Cassell G.H.;
RT   "The complete sequence of the mucosal pathogen Ureaplasma
RT   urealyticum.";
RL   Nature 407:757-762(2000).
DR   EMBL; AE002107; AAF30485.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   72 AA;  8562 MW;  5DFD611B9A9554B7 CRC64;

  Query Match             83.3%;  Score 5;  DB 16;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     50 RKLYD 54


RESULT 6
Q9X3B8
ID   Q9X3B8      PRELIMINARY;      PRT;    73 AA.
AC   Q9X3B8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ORF7.
OS   Pediococcus pentosaceus.
OG   Plasmid pMD136.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC   Pediococcus.
OX   NCBI_TaxID=1255;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC43200;
RA   Kantor A., Mett A., Shapira R.;
RT   "Pediococcus pentosaceus pediocin A encoding plasmid, pMD136.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC43200;
RA   Giacomini A., Squartini A., Nuti M.P.;
RT   "Nucleotide sequence and analysis of plasmid pMD136 from Pedi= ococcus
RT   pentosaceus FBB61 (ATCC43200) involved in Pediocin A production.";
RL   Plasmid 0:0-0(2000).
DR   EMBL; AF069302; AAD25913.1; -.
DR   EMBL; AF033858; AAF22854.1; -.
KW   Plasmid; Hypothetical protein.
SQ   SEQUENCE   73 AA;  8526 MW;  16277C452C104897 CRC64;

  Query Match             83.3%;  Score 5;  DB 2;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     47 RKLYD 51


RESULT 7
Q92HB1
ID   Q92HB1      PRELIMINARY;      PRT;    79 AA.
AC   Q92HB1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ANKYRIN LIKE PROTEIN.
GN   RC0860.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MALISH 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
DR   EMBL; AE008641; AAL03398.1; -.
KW   Complete proteome.
SQ   SEQUENCE   79 AA;  9340 MW;  FEA9F7581A040ED2 CRC64;

  Query Match             83.3%;  Score 5;  DB 16;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 66;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     21 RKLYD 25


RESULT 8
Q9PMK4
ID   Q9PMK4      PRELIMINARY;      PRT;   104 AA.
AC   Q9PMK4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE PERIPLASMIC PROTEIN.
GN   CJ1456C.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
```

```
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
DR   EMBL; AL139078; CAB73879.1; -.
KW   Complete proteome.
SQ   SEQUENCE   104 AA;  11853 MW;  BE26E94537C791E8 CRC64;

  Query Match             83.3%;  Score 5;  DB 16;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 82;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 KLYDY 6
           |||||
Db      53 KLYDY 57


RESULT 9
O52124
ID   O52124      PRELIMINARY;      PRT;   107 AA.
AC   O52124;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 12.3 KDA PROTEIN L0052 (ORF3).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98254123; PubMed=9593291;
RA   Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA   Lai L.C., McNamara B.P., Donnenberg M.S., Kaper J.B.;
RT   "The complete sequence of the locus of enterocyte effacement (LEE)
RT   from enteropathogenic Escherichia coli E2348/69.";
RL   Mol. Microbiol. 28:1-4(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL933;
RX   MEDLINE=98339885; PubMed=9673266;
RA   Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA   Kaper J.B., Blattner F.R.;
RT   "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT   Escherichia coli O157:H7.";
RL   Infect. Immun. 66:3810-3817(1998).
DR   EMBL; AF022236; AAC38366.1; -.
DR   EMBL; AF071034; AAC31531.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   107 AA;  12298 MW;  9B3FFBA7C8D448CE CRC64;

  Query Match             83.3%;  Score 5;  DB 2;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 84;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      68 RKLYD 72


RESULT 10
O32180
ID   O32180      PRELIMINARY;      PRT;   110 AA.
AC   O32180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   YUSN PROTEIN.
GN   YUSN.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99120; CAB15275.1; -.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   110 AA;  13138 MW;  17FE812FBE408253 CRC64;

  Query Match             83.3%;  Score 5;  DB 16;  Length 110;
  Best Local Similarity   100.0%;  Pred. No. 86;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      69 RKLYD 73


RESULT 11
Q9VAW7
ID   Q9VAW7      PRELIMINARY;      PRT;   111 AA.
AC   Q9VAW7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG17856 PROTEIN.
GN   CG17856.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
```

```
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003764; AAF56781.1; -.
DR   FlyBase; FBgn0039576; CG17856.
DR   InterPro; IPR003197; UCR_14kD.
DR   Pfam; PF02271; UCR_14kD; 1.
SQ   SEQUENCE   111 AA;  13517 MW;  839BE06F2E1B05D7 CRC64;

  Query Match             83.3%;  Score 5;  DB 5;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      52 RKLYD 56


RESULT 12
Q9VXI6
ID   Q9VXI6      PRELIMINARY;      PRT;   111 AA.
AC   Q9VXI6; Q9VXI5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG3560 PROTEIN.
GN   CG3560.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003501; AAF48576.1; -.
DR   FlyBase; FBgn0030733; CG3560.
DR   InterPro; IPR003197; UCR_14kD.
DR   Pfam; PF02271; UCR_14kD; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   111 AA;  13568 MW;  17E88E542F454624 CRC64;

  Query Match             83.3%;  Score 5;  DB 5;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYD 5
           |||||
Db      52 RKLYD 56


RESULT 13
Q9UNJ4
ID   Q9UNJ4      PRELIMINARY;      PRT;   129 AA.
AC   Q9UNJ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ESTROGEN RECEPTOR RELATED PROTEIN 3 (FRAGMENT).
GN   ERR3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
```

```
RX   MEDLINE=99357798; PubMed=10428842;
RA   Hong H., Yang L., Stallcup M.R.;
RT   "Hormone-independent transcriptional activation and coactivator
RT   binding by novel orphan nuclear receptor ERR3.";
RL   J. Biol. Chem. 274:22618-22626(1999).
DR   EMBL; AF117255; AAD48370.1; -.
KW   Receptor.
FT   NON_TER     129    129
SQ   SEQUENCE   129 AA;  13990 MW;  7BE487F716CCCD48 CRC64;

  Query Match             83.3%;  Score 5;  DB 4;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db     98 RKLYD 102


RESULT 14
Q98F36
ID   Q98F36      PRELIMINARY;      PRT;   148 AA.
AC   Q98F36;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   ACETYLTRANSFERASE.
GN   MLR3958.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003003; BAB50731.1; -.
DR   InterPro; IPR000182; Acetyltransf_GCN5.
DR   Pfam; PF00583; Acetyltransf; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   148 AA;  17075 MW;  A59FAEB203CBEADF CRC64;

  Query Match             83.3%;  Score 5;  DB 16;  Length 148;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYD 5
          |||||
Db    130 RKLYD 134


RESULT 15
Q9AUJ1
ID   Q9AUJ1      PRELIMINARY;      PRT;   157 AA.
AC   Q9AUJ1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MADS BOX PROTEIN NMADS1 (FRAGMENT).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yuan Z., Qian X., Liu J., Liu J., Qian M., Yang J., Liu J.;
RT   "Cloning and characterization of two cDNAs encoding rice MADS box
RT   protein.";
RL   Prog. Nat. Sci. 10:357-363(2000).
DR   EMBL; AF095645; AAK26240.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; K-box.
DR   InterPro; IPR002100; MADS-box.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   157 AA;  18145 MW;  05622210C95B4F32 CRC64;

  Query Match             83.3%;  Score 5;  DB 10;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLYDY 6
          |||||
Db     46 KLYDY 50


Search completed: November  8, 2002, 09:38:33
Job time : 27 secs
```

(11)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:21 ; Search time 4.66667 Seconds
(without alignments)
49.782 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 34
Sequence: 1 RKLYDY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	810	1	PLMN_HUMAN	P00747 homo sapien
2	34	100.0	812	1	PLMN_MOUSE	P20918 mus musculu
3	32	94.1	295	1	YFL1_YEAST	P43613 saccharomyc
4	32	94.1	439	1	GFO_ZYMMO	Q07982 zymomonas m
5	31	91.2	190	1	RL5_METJA	P54040 methanococc
6	31	91.2	434	1	ENO_STRIT	Q9xds7 streptococc
7	31	91.2	434	1	ENO_STRPY	P82479 streptococc
8	31	91.2	1113	1	Y140_MYCPN	P75033 mycoplasma
9	31	91.2	1161	1	RPO2_FOWPV	Q9j544 fowlpox vir
10	31	91.2	1534	1	DNM1_ARATH	P34881 arabidopsis
11	30	88.2	144	1	G10_XENLA	P12805 xenopus lae
12	30	88.2	333	1	PLMN_CANFA	P80009 canis famil
13	30	88.2	343	1	PLMN_SHEEP	P81286 ovis aries
14	30	88.2	426	1	YK96_AERPE	Q9ya42 aeropyrum p
15	30	88.2	504	1	SIK1_YEAST	Q12460 saccharomyc
16	30	88.2	810	1	PLMN_ERIEU	Q29485 erinaceus e
17	30	88.2	810	1	PLMN_MACMU	P12545 macaca mula
18	30	88.2	959	1	K6P2_YEAST	P16862 saccharomyc
19	30	88.2	4548	1	APOA_HUMAN	P08519 homo sapien
20	29	85.3	177	1	YA16_SCHPO	Q09688 schizosacch
21	29	85.3	305	1	DDL_BACHD	Q9kcf0 bacillus ha
22	29	85.3	405	1	DCDA_HELPJ	Q9zme5 helicobacte
23	29	85.3	405	1	DCDA_HELPY	P56129 helicobacte
24	29	85.3	488	1	EXON_HSV6U	P24447 human herpe
25	29	85.3	488	1	EXON_HSV6Z	P52448 human herpe
26	29	85.3	690	1	Y173_UREPA	Q9pqx0 ureaplasma
27	29	85.3	822	1	DEXT_STRSL	Q59979 streptococc
28	29	85.3	974	1	CC15_YEAST	P27636 saccharomyc
29	29	85.3	1202	1	RPM2_YEAST	Q02773 saccharomyc
30	28	82.4	90	1	MS2B_DROMA	P33738 drosophila
31	28	82.4	90	1	MS2B_DROME	P10334 drosophila
32	28	82.4	90	1	MS2B_DROSE	P33739 drosophila
33	28	82.4	90	1	MS2B_DROSI	P33740 drosophila
34	28	82.4	187	1	DDP1_YEAST	Q99321 saccharomyc
35	28	82.4	205	1	ENGB_BUCAI	P57507 buchnera ap
36	28	82.4	327	1	VP35_FOWPV	Q9j590 fowlpox vir
37	28	82.4	413	1	YV4Q_CAEEL	O45435 caenorhabdi
38	28	82.4	443	1	ADA2_HUMAN	O75478 homo sapien
39	28	82.4	505	1	VGLC_HSVMG	P10681 marek's dis
40	28	82.4	536	1	SCD2_SCHPO	P40996 schizosacch
41	28	82.4	724	1	P85A_BOVIN	P23727 bos taurus
42	28	82.4	724	1	P85A_HUMAN	P27986 homo sapien
43	28	82.4	724	1	P85A_MOUSE	P26450 mus musculu
44	28	82.4	724	1	P85A_RAT	Q63787 rattus norv
45	27	79.4	103	1	YQBF_BACSU	P45922 bacillus su

ALIGNMENTS

RESULT 1
PLMN_HUMAN
ID PLMN_HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;

RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vali Z., Patthy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vali Z., Patthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RT are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RT plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE 268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevich R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;
RL Submitted (JUN-1995) to the PDB data bank.
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.
RX MEDLINE=91095410; PubMed=1986355;
RA Ichinose A., Espling E.S., Takamatsu J., Saito H., Shinmyozu K.,

```
  Query Match             100.0%;  Score 34;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 2
PLMN_MOUSE
ID   PLMN_MOUSE     STANDARD;      PRT;   812 AA.
AC   P20918;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN   PLG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91184812; PubMed=2081600;
RA   Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse plasminogen and
RT   localization of the gene to mouse chromosome 17.";
RL   Genomics 8:49-61(1990).
RN   [2]
RP   CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX   MEDLINE=95042728; PubMed=7525077;
RA   O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA   Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT   "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT   suppression of metastases by a Lewis lung carcinoma.";
RL   Cell 79:315-328(1994).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC       NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC       METASTATIC TUMORS IN VIVO.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04766; AAA50168.1; -.
DR   PIR; A38514; A38514.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   MGD; MGI:97620; Plg.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
FT   CHAIN        20    581       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE      20     97       ACTIVATION PEPTIDE.
FT   CHAIN        98    581       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN        98   ?436       ANGIOSTATIN.
FT   CHAIN       582    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN      103    181       KRINGLE 1.
FT   DOMAIN      184    262       KRINGLE 2.
FT   DOMAIN      275    352       KRINGLE 3.
FT   DOMAIN      377    454       KRINGLE 4.
FT   DOMAIN      481    560       KRINGLE 5.
FT   DOMAIN      582    812       SERINE PROTEASE.
FT   ACT_SITE    624    624       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    667    667       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    762    762       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     49     73       BY SIMILARITY.
FT   DISULFID     53     61       BY SIMILARITY.
FT   DISULFID    103    181       BY SIMILARITY.
FT   DISULFID    124    164       BY SIMILARITY.
FT   DISULFID    152    176       BY SIMILARITY.
FT   DISULFID    185    262       BY SIMILARITY.
FT   DISULFID    188    316       BY SIMILARITY.
FT   DISULFID    206    245       BY SIMILARITY.
FT   DISULFID    234    257       BY SIMILARITY.
FT   DISULFID    275    352       BY SIMILARITY.
FT   DISULFID    296    335       BY SIMILARITY.
FT   DISULFID    324    347       BY SIMILARITY.
FT   DISULFID    377    454       BY SIMILARITY.
FT   DISULFID    398    437       BY SIMILARITY.
FT   DISULFID    426    449       BY SIMILARITY.
FT   DISULFID    481    560       BY SIMILARITY.
FT   DISULFID    502    543       BY SIMILARITY.
FT   DISULFID    531    555       BY SIMILARITY.
FT   DISULFID    568    687       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    578    586       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    609    625       BY SIMILARITY.
FT   DISULFID    701    768       BY SIMILARITY.
FT   DISULFID    731    747       BY SIMILARITY.
FT   DISULFID    758    786       BY SIMILARITY.
SQ   SEQUENCE   812 AA;  90846 MW;  D34A74A4FC2256F8 CRC64;

  Query Match             100.0%;  Score 34;  DB 1;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554
```

```
RESULT 3
YFL1_YEAST
ID   YFL1_YEAST     STANDARD;      PRT;   295 AA.
AC   P43613;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 34.2 kDa protein in SAP155-YMR31 intergenic region
DE   precursor.
GN   YFR041C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE=95400292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE=96287652; PubMed=8686379;
RA   Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA   Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT   "Analysis of a 36.2 kb DNA sequence including the right telomere of
RT   chromosome VI from Saccharomyces cerevisiae.";
RL   Yeast 12:149-167(1996).
CC   -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D50617; BAA09280.1; -.
DR   EMBL; D44597; BAA08007.1; -.
DR   HSSP; P25685; 1HDJ.
DR   SGD; S0001937; YFR041C.
DR   InterPro; IPR001623; DnaJ_N.
DR   Pfam; PF00226; DnaJ; 1.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS00636; DNAJ_1; 1.
DR   PROSITE; PS50076; DNAJ_2; 1.
KW   Hypothetical protein; Chaperone; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    295       HYPOTHETICAL PROTEIN YFR041C.
FT   DOMAIN       42    110       J-DOMAIN.
SQ   SEQUENCE   295 AA;  34191 MW;  4541BF6C353DB8AC CRC64;

  Query Match             94.1%;  Score 32;  DB 1;  Length 295;
  Best Local Similarity   83.3%;  Pred. No. 15;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||:|||
Db     101 RKIYDY 106


RESULT 4
GFO_ZYMMO
ID   GFO_ZYMMO      STANDARD;      PRT;   439 AA.
AC   Q07982;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucose--fructose oxidoreductase precursor (EC 1.1.99.28) (GFOR).
GN   GFO.
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC   Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 53-87.
RC   STRAIN=ATCC 29191 / ZM6;
RX   MEDLINE=92165717; PubMed=1537789;
RA   Kanagasundaram V., Scopes R.K.;
RT   "Cloning, sequence analysis, and expression of the structural gene
RT   encoding glucose-fructose oxidoreductase from Zymomonas mobilis.";
RL   J. Bacteriol. 174:1439-1447(1992).
RN   [2]
RP   SEQUENCE OF 1-71 FROM N.A., AND SEQUENCE OF 2-16 AND 53-71.
RC   STRAIN=ATCC 29191 / ZM6, AND ATCC 31821 / ZM4 / CP4;
RX   MEDLINE=93231476; PubMed=8472911;
RA   Loos H., Sahm H., Sprenger G.A.;
RT   "Glucose-fructose oxidoreductase, a periplasmic enzyme of Zymomonas
RT   mobilis, is active in its precursor form.";
RL   FEMS Microbiol. Lett. 107:293-298(1993).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX   MEDLINE=97148336; PubMed=8994968;
RA   Kingston R.L., Scopes R.K., Baker E.N.;
RT   "The structure of glucose-fructose oxidoreductase from Zymomonas
RT   mobilis: an osmoprotective periplasmic enzyme containing non-
RT   dissociable NADP.";
RL   Structure 4:1413-1428(1996).
CC   -!- CATALYTIC ACTIVITY: D-glucose + D-fructose = D-gluconolactone + D-
CC       glucitol.
CC   -!- COFACTOR: CONTAINS ONE TIGHTLY-BOUND NADP(H) MOLECULE PER SUBUNIT.
CC   -!- PATHWAY: SORBITOL-GLUCONATE PATHWAY.
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M97379; AAA27690.1; -.
DR   EMBL; X73088; CAA51534.1; -.
DR   PDB; 1OFG; 21-APR-97.
DR   InterPro; IPR000683; GFO_IDH_MocA.
DR   InterPro; IPR004104; GFO_IDH_MocA_C.
DR   Pfam; PF01408; GFO_IDH_MocA; 1.
DR   Pfam; PF02894; GFO_IDH_MocA_C; 1.
KW   Oxidoreductase; NADP; Signal; Periplasmic; 3D-structure.
FT   SIGNAL        1     52
FT   CHAIN        53    439       GLUCOSE--FRUCTOSE OXIDOREDUCTASE.
FT   CONFLICT     61     61       L -> Q (IN REF. 2; AA SEQUENCE).
SQ   SEQUENCE   439 AA;  47801 MW;  C1E4A19F5B34267A CRC64;

  Query Match             94.1%;  Score 32;  DB 1;  Length 439;
  Best Local Similarity   83.3%;  Pred. No. 22;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||:|||
Db     133 RKIYDY 138


RESULT 5
RL5_METJA
ID   RL5_METJA      STANDARD;      PRT;   190 AA.
AC   P54040;
```

```
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L5P.
GN   RPL5P OR MJ0469.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U67497; AAB98458.1; -.
DR   TIGR; MJ0469; -.
DR   InterPro; IPR002132; Ribosomal_L5.
DR   Pfam; PF00281; Ribosomal_L5; 1.
DR   Pfam; PF00673; Ribosomal_L5_C; 1.
DR   ProDom; PD001076; Ribosomal_L5; 1.
DR   PROSITE; PS00358; RIBOSOMAL_L5; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   190 AA;  22169 MW;  BD05366973C3A533 CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 190;
  Best Local Similarity   83.3%;  Pred. No. 16;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
             :|||||
Db      95 KKLYDY 100


RESULT 6
ENO_STRIT
ID   ENO_STRIT      STANDARD;      PRT;   434 AA.
AC   Q9XDS7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE   glycerate hydro-lyase).
GN   ENO.
OS   Streptococcus intermedius.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1338;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 27335;
RA   Sato S.;
RT   "Streptococcus intermedius enolase gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC       H(2)O.
CC   -!- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC       THE DIMER (BY SIMILARITY).
CC   -!- PATHWAY: GLYCOLYSIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB029313; BAA81815.1; -.
DR   HSSP; P00924; 4ENL.
DR   InterPro; IPR000941; Enolase.
DR   Pfam; PF00113; enolase; 1.
DR   PRINTS; PR00148; ENOLASE.
DR   ProDom; PD000902; Enolase; 1.
DR   PROSITE; PS00164; ENOLASE; 1.
KW   Lyase; Glycolysis; Magnesium.
FT   ACT_SITE    155    155       BY SIMILARITY.
FT   METAL       242    242       MAGNESIUM (BY SIMILARITY).
FT   METAL       291    291       MAGNESIUM (BY SIMILARITY).
FT   METAL       318    318       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   434 AA;  47015 MW;  AA0D39DA832B4F3B CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 434;
  Best Local Similarity   83.3%;  Pred. No. 35;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
             ||:|||
Db     253 RKVYDY 258


RESULT 7
ENO_STRPY
ID   ENO_STRPY      STANDARD;      PRT;   434 AA.
AC   P82479;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE   glycerate hydro-lyase).
GN   ENO OR SPY0731.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN   [2]
RP   PARTIAL SEQUENCE, AND MASS SPECTROMETRY.
RC   STRAIN=JRS4 / Serotype M6;
RA   Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA   VanBogelen R.A.;
RT   "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT   proteins.";
RL   Submitted (MAY-2000) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
```

```
CC       H(2)O.
CC   -!- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC       THE DIMER (BY SIMILARITY).
CC   -!- PATHWAY: GLYCOLYSIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006524; AAK33680.1; -.
DR   HSSP; P00924; 1ONE.
DR   InterPro; IPR000941; Enolase.
DR   Pfam; PF00113; enolase; 4.
DR   PRINTS; PR00148; ENOLASE.
DR   ProDom; PD000902; Enolase; 1.
DR   PROSITE; PS00164; ENOLASE; 1.
KW   Lyase; Glycolysis; Magnesium; Complete proteome.
FT   INIT_MET      0      0
FT   ACT_SITE    154    154       BY SIMILARITY.
FT   METAL       242    242       MAGNESIUM (BY SIMILARITY).
FT   METAL       291    291       MAGNESIUM (BY SIMILARITY).
FT   METAL       318    318       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   434 AA;  47225 MW;  C4FDAFC4D7D25B44 CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 434;
  Best Local Similarity   83.3%;  Pred. No. 35;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    253 RKVYDY 258


RESULT 8
Y140_MYCPN
ID   Y140_MYCPN     STANDARD;      PRT;  1113 AA.
AC   P75033;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical ATP-binding protein MG140 homolog (E07_orf1113).
GN   MPN153 OR MP001.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SIMILARITY: BELONGS TO THE DNA2/NAM7 HELICASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000001; AAB95649.1; -.
KW   Hypothetical protein; ATP-binding; Helicase; Complete proteome.
FT   NP_BIND     313    320       ATP (POTENTIAL).
SQ   SEQUENCE   1113 AA;  130333 MW;  48A3337EB0E81A40 CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 1113;
  Best Local Similarity   83.3%;  Pred. No. 90;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          |:||||
Db    693 RRLYDY 698


RESULT 9
RPO2_FOWPV
ID   RPO2_FOWPV     STANDARD;      PRT;  1161 AA.
AC   Q9J544;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN   RPO132 OR FPV189.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC   -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44533.1; -.
DR   InterPro; IPR001572; RNA_pol_B.
DR   Pfam; PF00562; RNA_pol_B; 1.
DR   PROSITE; PS01166; RNA_POL_BETA; 1.
KW   Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW   Zinc-finger.
FT   ZN_FING    1079   1098       C4-TYPE (POTENTIAL).
SQ   SEQUENCE   1161 AA;  133468 MW;  334C2FC5DD63584F CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 1161;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          :|||||
Db    647 KKLYDY 652


RESULT 10
DNM1_ARATH
ID   DNM1_ARATH     STANDARD;      PRT;  1534 AA.
AC   P34881;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
```

```
DE   DNA (cytosine-5)-methyltransferase AthI (EC 2.1.1.37) (DNA
DE   methyltransferase AthI) (DNA Metase AthI) (M.AthI).
GN   ATHIM OR AT5G49160 OR K21P3.3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=93281384; PubMed=8389441;
RA   Finnegan E.J., Dennis E.S.;
RT   "Isolation and identification by sequence homology of a putative
RT   cytosine methyltransferase from Arabidopsis thaliana.";
RL   Nucleic Acids Res. 21:2383-2388(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=99156233; PubMed=10048488;
RA   Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
RT   Sequence features of the regions of 1,081,958 bp covered by seventeen
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:379-391(1998).
CC   -!- FUNCTION: METHYLATES CG RESIDUES.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA = S-adenosyl-L-
CC       homocysteine + DNA containing 5-methylcytosine.
CC   -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 BAH DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10692; AAA32829.1; -.
DR   EMBL; AB016872; BAB10334.1; -.
DR   REBASE; 2839; M.AthI.
DR   InterPro; IPR001025; BAH.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   Pfam; PF01426; BAH; 2.
DR   Pfam; PF00145; DNA_methylase; 3.
DR   PRINTS; PR00105; C5METTRFRASE.
DR   SMART; SM00439; BAH; 2.
DR   PROSITE; PS00094; C5_MTASE_1; 1.
DR   PROSITE; PS00095; C5_MTASE_2; 1.
KW   Transferase; Methyltransferase; DNA-binding.
FT   ACT_SITE   1198   1198       BY SIMILARITY.
SQ   SEQUENCE   1534 AA;  172430 MW;  23FC944AA7074C5A CRC64;

  Query Match             91.2%;  Score 31;  DB 1;  Length 1534;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    172 RKVYDY 177


RESULT 11
G10_XENLA
ID   G10_XENLA      STANDARD;      PRT;   144 AA.
AC   P12805;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   G10 protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89306617; PubMed=2568313;
RA   McGrew L.L., Dworkin-Rastl E., Dworkin M.B., Richter J.D.;
RT   "Poly(A) elongation during Xenopus oocyte maturation is required for
RT   translational recruitment and is mediated by a short sequence
RT   element.";
RL   Genes Dev. 3:803-815(1989).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- DEVELOPMENTAL STAGE: OOCYTE MATURATION IS ACCOMPANIED BY THE
CC       RECRUITMENT OF SPECIFIC MATERNAL MRNAS INTO POLYSOMES. G10 IS AN
CC       EXAMPLE OF A PROTEIN WHICH IS TRANSLATED AT THAT TIME.
CC   -!- SIMILARITY: BELONGS TO THE G10 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15243; CAA33321.1; -.
DR   PIR; S05955; S05955.
DR   InterPro; IPR001748; G10.
DR   Pfam; PF01125; G10; 1.
DR   PRINTS; PR00322; G10.
DR   ProDom; PD009460; G10; 1.
DR   PROSITE; PS00997; G10_1; 1.
DR   PROSITE; PS00998; G10_2; 1.
KW   Developmental protein; Nuclear protein; Zinc-finger.
FT   DOMAIN        2     10       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   ZN_FING     101    119       POTENTIAL.
SQ   SEQUENCE   144 AA;  17040 MW;  D2A7D3D4E398B696 CRC64;

  Query Match             88.2%;  Score 30;  DB 1;  Length 144;
  Best Local Similarity   83.3%;  Pred. No. 19;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          |:||||
Db     72 RELYDY 77


RESULT 12
PLMN_CANFA
ID   PLMN_CANFA     STANDARD;      PRT;   333 AA.
AC   P80009;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=90175323; PubMed=2626424;
RA   Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT   "Complete amino acid sequence of canine miniplasminogen.";
RL   Protein Seq. Data Anal. 2:445-450(1989).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
```

```
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF
CC       STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   CHAIN        <1    103       PLASMIN HEAVY CHAIN A.
FT   CHAIN       104    333       PLASMIN LIGHT CHAIN B.
FT   DOMAIN        4     83       KRINGLE 5.
FT   DOMAIN      104    333       SERINE PROTEASE.
FT   DISULFID      4     83       BY SIMILARITY.
FT   DISULFID     25     66       BY SIMILARITY.
FT   DISULFID     54     78       BY SIMILARITY.
FT   DISULFID     90    208       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    100    108       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    130    146       BY SIMILARITY.
FT   DISULFID    222    289       BY SIMILARITY.
FT   DISULFID    252    268       BY SIMILARITY.
FT   DISULFID    279    307       BY SIMILARITY.
FT   ACT_SITE    145    145       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    188    188       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    283    283       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE        152    152       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        186    186       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        264    264       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        277    277       SITE OF SUBSTRATE SPECIFICITY
FT                                 (BY SIMILARITY).
SQ   SEQUENCE   333 AA;  36678 MW;  C8C0271B6C6AC8D4 CRC64;

  Query Match             88.2%;  Score 30;  DB 1;  Length 333;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           |||:||
Db      72 RKLFDY 77


RESULT 13
PLMN_SHEEP
ID   PLMN_SHEEP     STANDARD;      PRT;   343 AA.
AC   P81286;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93149995; PubMed=1492092;
RA   Schaller J., Straub C., Kampfer U., Rickli E.E.;
RT   "Complete amino acid sequence of ovine miniplasminogen.";
RL   Protein Seq. Data Anal. 5:21-25(1992).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 2 KRINGLE DOMAINS.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   DOMAIN       <1    140       HEAVY CHAIN A.
FT   DOMAIN      141   >343       LIGHT CHAIN A.
FT   DOMAIN       <1     17       KRINGLE 4.
FT   DOMAIN       41    120       KRINGLE 5.
FT   DOMAIN      114    341       SERINE PROTEASE.
FT   ACT_SITE    181    181       CHARGE RELAY SYSTEM.
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM.
FT   ACT_SITE    319    319       CHARGE RELAY SYSTEM.
FT   NON_TER     343    343
SQ   SEQUENCE   343 AA;  37662 MW;  8DF6EBA92D596EE0 CRC64;

  Query Match             88.2%;  Score 30;  DB 1;  Length 343;
  Best Local Similarity   83.3%;  Pred. No. 45;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           |||:||
Db      83 RKLFDY 88


RESULT 14
YK96_AERPE
ID   YK96_AERPE     STANDARD;      PRT;   426 AA.
AC   Q9YA42;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein APE2096.
GN   APE2096.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
CC   -!- SIMILARITY: BELONGS TO THE UPF0095 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP000063; BAA81107.1; -.
DR   InterPro; IPR002936; Toprim.
DR   Pfam; PF01751; Toprim; 1.
DR   SMART; SM00493; TOPRIM; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   426 AA;  47681 MW;  73A0B78FCD1206B1 CRC64;

  Query Match             88.2%;  Score 30;  DB 1;  Length 426;
  Best Local Similarity   83.3%;  Pred. No. 56;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           |:||||
Db     362 RELYDY 367


RESULT 15
SIK1_YEAST
ID   SIK1_YEAST     STANDARD;      PRT;   504 AA.
AC   Q12460;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   SIK1 protein (Nucleolar protein NOP56).
GN   SIK1 OR NOP56 OR YLR197W OR L8167.9.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YPH1;
RX   MEDLINE=96040178; PubMed=7547500;
RA   Morin P.J., Downs J.A., Snodgrass A.M., Gilmore T.D.;
RT   "Genetic analysis of growth inhibition by GAL4-L kappa B-alpha in
RT   Saccharomyces cerevisiae.";
RL   Cell Growth Differ. 6:789-798(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA   Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA   Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA   Johnston L., Langston Y., Latreille P., Mardis E., Menezes S.,
RA   Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA   Taich S., Trevaskis E., Vaudin M., Vignati D., Wilcox L., Wilson R.,
RA   Wohldman P., Waterston R.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   CHARACTERIZATION, AND MUTAGENESIS.
RX   MEDLINE=98038777; PubMed=9372940;
RA   Gautier T., Berges T., Tollervey D., Hurt E.;
RT   "Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p
RT   and are required for ribosome biogenesis.";
RL   Mol. Cell. Biol. 17:7088-7098(1997).
CC   -!- FUNCTION: REQUIRED FOR 60S RIBOSOMAL SUBUNIT SYNTHESIS.
CC   -!- SUBUNIT: INTERACTS WITH NOP1 AND NOP58.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U20237; AAC49066.1; -.
DR   EMBL; U14913; AAB67431.1; -.
DR   SGD; S0004187; SIK1.
DR   InterPro; IPR002687; Nop.
DR   Pfam; PF01798; Nop; 1.
DR   ProDom; PD004104; Nop; 1.
KW   Ribosome biogenesis; Nuclear protein.
FT   DOMAIN      443    504       ASP/GLU/LYS-RICH.
FT   MUTAGEN     333    333       V->A: REDUCED GROWTH RATE AT ALL
FT                                TEMPERATURES; WHEN ASSOCIATED WITH R-385.
FT   MUTAGEN     355    355       Y->C: AT 37 DEGREES, GROWTH SLOWS AFTER 6
FT                                TO 8 HOURS AND CELL DIVISION STOPS AFTER
FT                                20 HOURS.
FT   MUTAGEN     385    385       M->R: REDUCED GROWTH RATE AT ALL
FT                                TEMPERATURES; WHEN ASSOCIATED WITH A-333.
SQ   SEQUENCE   504 AA;  56864 MW;  F8522A5870EF4842 CRC64;

  Query Match             88.2%;  Score 30;  DB 1;  Length 504;
  Best Local Similarity   83.3%;  Pred. No. 66;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           |:||||
Db     286 RQLYDY 291


Search completed: November  8, 2002, 09:33:54
Job time : 6.66667 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:22 ; Search time 14.3333 Seconds
(without alignments)
72.416 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 34
Sequence: 1 RKLYDY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	334	6	O46507	O46507 papio hamad
2	34	100.0	810	4	Q15146	Q15146 homo sapien
3	34	100.0	812	11	Q9R0W3	Q9r0w3 rattus norv
4	34	100.0	812	11	Q91WJ5	Q91wj5 mus musculu
5	32	94.1	73	12	Q07273	Q07273 turkey rhin
6	32	94.1	254	16	Q9PMJ3	Q9pmj3 campylobact
7	32	94.1	404	16	Q9KVG1	Q9kvg1 vibrio chol
8	32	94.1	433	2	P75002	P75002 zymomonas m
9	31	91.2	274	5	P91014	P91014 caenorhabdi
10	31	91.2	388	9	Q38198	Q38198 xanthomonas
11	31	91.2	409	16	Q99TG1	Q99tg1 staphylococ
12	31	91.2	417	2	Q45037	Q45037 borrelia bu
13	31	91.2	417	16	O50899	O50899 borrelia bu
14	31	91.2	434	2	Q935W7	Q935w7 streptococc
15	31	91.2	434	16	Q97QS2	Q97qs2 streptococc
16	31	91.2	454	10	Q9ZRW0	Q9zrw0 cicer ariet
17	31	91.2	499	10	Q9LTV0	Q9ltv0 arabidopsis
18	31	91.2	522	10	Q9SGT7	Q9sgt7 arabidopsis
19	31	91.2	584	5	Q9U0M8	Q9u0m8 plasmodium
20	31	91.2	633	5	O97305	O97305 plasmodium
21	31	91.2	1165	12	Q98839	Q98839 molluscum c
22	31	91.2	1404	10	Q9T0I1	Q9t0i1 arabidopsis
23	31	91.2	1517	10	Q9SEG3	Q9seg3 arabidopsis
24	31	91.2	1519	10	O23273	O23273 arabidopsis
25	30	88.2	144	5	O97454	O97454 drosophila
26	30	88.2	160	17	Q97CJ1	Q97cj1 thermoplasm
27	30	88.2	292	16	Q97QJ3	Q97qj3 streptococc
28	30	88.2	294	16	Q99ZF5	Q99zf5 streptococc
29	30	88.2	319	16	Q97DR3	Q97dr3 clostridium
30	30	88.2	320	5	Q18887	Q18887 caenorhabdi
31	30	88.2	405	9	O64073	O64073 bacteriopha
32	30	88.2	405	16	O31948	O31948 bacillus su
33	30	88.2	474	17	O28424	O28424 archaeoglob
34	30	88.2	502	16	Q9PQW1	Q9pqw1 ureaplasma
35	30	88.2	647	5	O44560	O44560 caenorhabdi
36	30	88.2	722	5	Q9V5F2	Q9v5f2 drosophila
37	30	88.2	732	2	Q9L281	Q9l281 streptomyce
38	30	88.2	1187	10	Q9ZV43	Q9zv43 arabidopsis
39	30	88.2	1436	3	Q07527	Q07527 saccharomyc
40	29	85.3	73	12	Q90089	Q90089 turkey rhin
41	29	85.3	104	16	Q9PMK4	Q9pmk4 campylobact
42	29	85.3	146	11	Q924Q3	Q924q3 mus musculu
43	29	85.3	148	16	Q98F36	Q98f36 rhizobium l
44	29	85.3	157	10	Q9AUJ1	Q9auj1 oryza sativ
45	29	85.3	170	2	Q9KX31	Q9kx31 streptococc

ALIGNMENTS

RESULT 1
O46507
ID O46507 PRELIMINARY; PRT; 334 AA.
AC O46507;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PLASMINOGEN (FRAGMENT).
GN BABPEPSG.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF029692; AAB97887.1; -.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.

```
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  36791 MW;  C7DC06E03B965286 CRC64;

  Query Match             100.0%;  Score 34;  DB 6;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db     73 RKLYDY 78


RESULT 2
Q15146
ID   Q15146      PRELIMINARY;      PRT;   810 AA.
AC   Q15146;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA   Mitchell D., Robinson J.H.;
RT   "Expression of recombinant human plasminogen and aglycoplasminogen in
RT   HeLa cells.";
RL   Fibrinolysis 0:0-0(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; M74220; AAA36451.1; -.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    810       PLASMINOGEN.
SQ   SEQUENCE   810 AA;  90555 MW;  B05C7D4B0D020B3C CRC64;

  Query Match             100.0%;  Score 34;  DB 4;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 3
Q9R0W3
ID   Q9R0W3      PRELIMINARY;      PRT;   812 AA.
AC   Q9R0W3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PROTEIN PRECURSOR (EC 3.4.21.7).
GN   PLASMINOGEN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Bangert K., Johnsen A.H., Thorsen S.;
RT   "Rat plasminogen: cDNA and gene structure.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=91250378; PubMed=1645711;
RA   Kanalas J.J., Makker S.P.;
RT   "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT   receptor site for plasminogen.";
RL   J. Biol. Chem. 266:10825-10829(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AJ242649; CAB46014.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
SQ   SEQUENCE   812 AA;  90535 MW;  8C703C51410EBC9E CRC64;

  Query Match             100.0%;  Score 34;  DB 11;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 4
Q91WJ5
ID   Q91WJ5      PRELIMINARY;      PRT;   812 AA.
AC   Q91WJ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014773; AAH14773.1; -.
SQ   SEQUENCE   812 AA;  90781 MW;  24173260E6A2FFD2 CRC64;

  Query Match             100.0%;  Score 34;  DB 11;  Length 812;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db    549 RKLYDY 554


RESULT 5
Q07273
ID   Q07273      PRELIMINARY;      PRT;    73 AA.
AC   Q07273;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-FEB-1997 (TrEMBLrel. 02, Last annotation update)
DE   HYPOTHETICAL 8.5 KDA PROTEIN IN 22K 3'REGION.
OS   Turkey rhinotracheitis virus (TRTV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX   NCBI_TaxID=11264;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UK/3BV/85;
RA   Yu Q., Davis P.J., Brown T.D.K., Cavanagh D.;
RL   J. Gen. Virol. 73:1355-1363(1993).
DR   EMBL; X63408; CAA45005.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   73 AA;  8488 MW;  D6E09B8BDC40EC4F CRC64;

  Query Match             94.1%;  Score 32;  DB 12;  Length 73;
  Best Local Similarity   83.3%;  Pred. No. 13;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db     25 RKIYDY 30


RESULT 6
Q9PMJ3
ID   Q9PMJ3      PRELIMINARY;      PRT;   254 AA.
AC   Q9PMJ3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN CJ1467.
GN   CJ1467.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
DR   EMBL; AL139078; CAB73890.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   254 AA;  30201 MW;  15712EC3A70AD3E1 CRC64;

  Query Match             94.1%;  Score 32;  DB 16;  Length 254;
  Best Local Similarity   83.3%;  Pred. No. 48;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    210 RKIYDY 215


RESULT 7
Q9KVG1
ID   Q9KVG1      PRELIMINARY;      PRT;   404 AA.
AC   Q9KVG1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TRANSPOSASE, PUTATIVE.
GN   VC0185.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004108; AAF93361.1; -.
DR   TIGR; VC0185; -.
KW   Complete proteome.
SQ   SEQUENCE   404 AA;  47642 MW;  26A22EABFCBB4EBD CRC64;

  Query Match             94.1%;  Score 32;  DB 16;  Length 404;
  Best Local Similarity   83.3%;  Pred. No. 78;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    273 RKMYDY 278


RESULT 8
P75002
ID   P75002      PRELIMINARY;      PRT;   433 AA.
AC   P75002;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLUCOSE-FRUCTOSE OXIDOREDUCTASE PRECURSOR.
GN   GFO.
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC   Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZM6;
RA   Wiegert T., Sahm H., Sprenger G.A.;
RT   "The substitution of a single amino acid residue alters NADP-
```

```
RT   containing glucose-fructose oxidoreductase of Zymomonas mobilis . . .";
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z80356; CAB02496.1; -.
DR   HSSP; Q07982; 1OFG.
DR   InterPro; IPR000683; GFO_IDH_MocA.
DR   InterPro; IPR004104; GFO_IDH_MocA_C.
DR   Pfam; PF01408; GFO_IDH_MocA; 1.
DR   Pfam; PF02894; GFO_IDH_MocA_C; 1.
KW   Signal.
FT   SIGNAL        1     52
FT   CHAIN        53    433
SQ   SEQUENCE   433 AA;  47189 MW;  13CFADE84794E736 CRC64;

  Query Match             94.1%;  Score 32;  DB 2;  Length 433;
  Best Local Similarity   83.3%;  Pred. No. 84;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    134 RKIYDY 139


RESULT 9
P91014
ID   P91014      PRELIMINARY;      PRT;   274 AA.
AC   P91014;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 32.6 KDA PROTEIN.
GN   C01G8.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Du Z., Gattung S.;
RT   "The sequence of C. elegans cosmid C01G8.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U80439; AAB37641.2; -.
DR   HSSP; P25685; 1HDJ.
DR   InterPro; IPR003095; DnaJ.
DR   InterPro; IPR001623; DnaJ_N.
DR   Pfam; PF00226; DnaJ; 1.
DR   PRINTS; PR00625; DNAJPROTEIN.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS50076; DNAJ_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   274 AA;  32579 MW;  8A6F3C7692A572E9 CRC64;

  Query Match             91.2%;  Score 31;  DB 5;  Length 274;
  Best Local Similarity   83.3%;  Pred. No. 85;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          |:||||
Db     85 RRLYDY 90


RESULT 10
Q38198
ID   Q38198      PRELIMINARY;      PRT;   388 AA.
AC   Q38198;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BACTERIOPHAGE CF1C GENE.
OS   Xanthomonas phage Cf1c.
OC   Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX   NCBI_TaxID=10862;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91252299; PubMed=1840658;
RA   Kuo T.T., Tan M.S., Su M.T., Yang M.K.;
RT   "Nucleotide sequence of filamentous phage Cf1c from Xanthomonas
RT   campestris pv. citri.";
RL   Nucleic Acids Res. 19:2498-2498(1991).
DR   EMBL; M57538; AAA32201.1; -.
SQ   SEQUENCE   388 AA;  42785 MW;  4ED8ECCE19C0B809 CRC64;

  Query Match             91.2%;  Score 31;  DB 9;  Length 388;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          :|||||
Db    155 KKLYDY 160


RESULT 11
Q99TG1
ID   Q99TG1      PRELIMINARY;      PRT;   409 AA.
AC   Q99TG1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SA1524 PROTEIN (MALATE DEHYDROGENASE HOMOLOG).
GN   SA1524 OR SAV1702.
OS   Staphylococcus aureus (strain N315), and
OS   Staphylococcus aureus (strain Mu50).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=158879, 158878;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003134; BAB42791.1; -.
DR   EMBL; AP003363; BAB57864.1; -.
DR   InterPro; IPR001891; Malic_enzyme.
DR   InterPro; IPR000205; NAD_binding.
DR   Pfam; PF00390; malic; 2.
DR   PRINTS; PR00072; MALOXRDTASE.
DR   PROSITE; PS00331; MALIC_ENZYMES; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   409 AA;  44251 MW;  1D88682BD01F1EBD CRC64;

  Query Match             91.2%;  Score 31;  DB 16;  Length 409;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 RKLYDY 6
          ||:|||
Db     52 RKVYDY 57


RESULT 12
Q45037
ID   Q45037      PRELIMINARY;      PRT;   417 AA.
AC   Q45037;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   S1 PROTEIN PRECURSOR.
GN   S1.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N40;
RX   MEDLINE=95369900; PubMed=7642278;
RA   Feng S., Das S., Lam T., Flavell R.A., Fikrig E.;
RT   "A 55-kilodalton antigen encoded by a gene on a Borrelia burgdorferi
RT   49-kilobase plasmid is recognized by antibodies in sera from patients
RT   with Lyme disease.";
RL   Infect. Immun. 63:3459-3466(1995).
DR   EMBL; L34016; AAA81351.1; -.
KW   Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    417       S1 PROTEIN.
SQ   SEQUENCE   417 AA;  48869 MW;  F691721D194B33D3 CRC64;

  Query Match             91.2%;  Score 31;  DB 2;  Length 417;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          :|||||
Db    397 KKLYDY 402


RESULT 13
O50899
ID   O50899      PRELIMINARY;      PRT;   417 AA.
AC   O50899;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ANTIGEN, S1.
GN   BBA05.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
DR   EMBL; AE000790; AAC66229.1; -.
DR   TIGR; BBA05; -.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   417 AA;  48841 MW;  ED6A0FE883A8DDBA CRC64;

  Query Match             91.2%;  Score 31;  DB 16;  Length 417;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          :|||||
Db    397 KKLYDY 402


RESULT 14
Q935W7
ID   Q935W7      PRELIMINARY;      PRT;   434 AA.
AC   Q935W7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-ENOLASE.
GN   ENO.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TYPE 2 ATCC 11733;
RA   Bergmann S., Rohde M., Chhatwal G.S., Hammerschmidt S.;
RT   "alpha-enolase of Streptococcus pneumoniae is a plasminogen-binding
RT   protein displayed on the bacterial cell surface.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ303085; CAC83091.1; -.
FT   CHAIN         2    434       ALPHA-ENOLASE.
SQ   SEQUENCE   434 AA;  47075 MW;  9298639E949F61CE CRC64;

  Query Match             91.2%;  Score 31;  DB 2;  Length 434;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||:|||
Db    253 RKVYDY 258


RESULT 15
Q97QS2
ID   Q97QS2      PRELIMINARY;      PRT;   434 AA.
AC   Q97QS2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ENOLASE.
GN   SP1128.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007414; AAK75238.1; -.
```

```
DR   TIGR; SP1128; -.
DR   InterPro; IPR000941; Enolase.
DR   Pfam; PF00113; enolase; 1.
DR   PRINTS; PR00148; ENOLASE.
DR   ProDom; PD000902; Enolase; 1.
DR   PROSITE; PS00164; ENOLASE; 1.
KW   Complete proteome.
SQ   SEQUENCE   434 AA;  47103 MW;  0D64F8F04BBB99C4 CRC64;

  Query Match            91.2%;  Score 31;  DB 16;  Length 434;
  Best Local Similarity  83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
             ||:|||
Db     253 RKVYDY 258


Search completed: November  8, 2002, 09:36:12
Job time : 16.3333 secs
```

(11)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:36 ; Search time 9 Seconds
(without alignments)
64.060 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 34
Sequence: 1 RKLYDY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	810	1	PLHU	plasmin (EC 3.4.21
2	34	100.0	812	1	PLMS	plasmin (EC 3.4.21
3	32	94.1	73	2	JQ1988	hypothetical 8.5K
4	32	94.1	254	2	B81293	hypothetical prote
5	32	94.1	295	2	S56296	probable membrane
6	32	94.1	404	2	D82355	probable transposa
7	32	94.1	439	2	A42289	glucose-fructose o
8	31	91.2	142	2	A97409	hypothetical prote
9	31	91.2	190	2	E64358	ribosomal protein
10	31	91.2	345	2	T29261	hypothetical prote
11	31	91.2	388	2	S26824	hypothetical prote
12	31	91.2	409	2	B89954	hypothetical prote
13	31	91.2	417	2	E70207	antigen S1 - Lyme
14	31	91.2	434	2	E95130	enolase [imported]
15	31	91.2	434	2	D98001	phosphopyruvate hy
16	31	91.2	522	2	D96602	nucleolar protein
17	31	91.2	1113	2	S73327	MG140 homolog - My
18	31	91.2	1165	2	T30731	probable DNA-direc
19	31	91.2	1404	2	T06663	DNA (cytosine-5-)-
20	31	91.2	1519	2	G71402	DNA (cytosine-5-)-
21	31	91.2	1534	2	S59604	DNA (cytosine-5-)-
22	30	88.2	144	2	S05955	G10 protein - Afri
23	30	88.2	292	2	F95140	tRNA pseudouridine
24	30	88.2	292	2	C98008	pseudouridylate sy
25	30	88.2	314	2	AF0901	tRNA pseudouridine
26	30	88.2	319	2	H97318	transcription regu
27	30	88.2	320	2	T15849	hypothetical prote
28	30	88.2	405	2	T12824	hypothetical prote
29	30	88.2	426	2	C72515	hypothetical prote
30	30	88.2	460	2	B61545	plasmin (EC 3.4.21
31	30	88.2	474	2	E69481	4-hydroxybutyrate
32	30	88.2	502	2	A82925	phosphoglycerate m
33	30	88.2	504	2	S48550	hypothetical prote
34	30	88.2	647	2	T30892	hypothetical prote
35	30	88.2	810	2	B30848	plasmin (EC 3.4.21
36	30	88.2	810	2	I46260	plasmin (EC 3.4.21
37	30	88.2	959	1	JQ0017	6-phosphofructokin
38	30	88.2	1187	2	C84568	hypothetical prote
39	30	88.2	1436	2	S67655	probable membrane
40	30	88.2	4548	1	S00657	apoprotein(a) (EC
41	29	85.3	73	2	JQ1624	hypothetical 8.5K
42	29	85.3	104	2	G81291	probable periplasm
43	29	85.3	117	2	AI2440	hypothetical prote
44	29	85.3	126	2	AE0091	probable flagellar
45	29	85.3	177	2	S59643	hypothetical prote

ALIGNMENTS

RESULT 1
PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N;Alternate names: plasminogen precursor [misnomer]
N;Contains: angiostatin; microplasmin; plasminogen
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C;Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the f
A;Reference number: A35229; MUID:90202879
A;Accession: A35229
A;Molecule type: DNA
A;Residues: 1-810 <PET>
A;Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A;Experimental source: leukocyte; lung fibroblast
R;Malgaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A;Title: Definition of the transcription initiation site of human plasminogen gene in
A;Reference number: I52242; MUID:91097523
A;Accession: I52242
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <MAL1>
A;Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613
R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human
A;Reference number: A26646; MUID:87162490
A;Accession: A26646
A;Molecule type: mRNA
A;Residues: 1-471,'D',473-810 <FOR>
A;Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
A;Experimental source: liver
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human a
A;Reference number: I45961; MUID:85023311
A;Accession: I62738
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 292-471,'D',473-810 <MAL2>
A;Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A;Accession: I84609
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 367-419 <MAL3>
A;Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg,
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of huma

A;Reference number: S03735; MUID:81212097
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71,'E',73-76 <BRU>
R;Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: protein
A;Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R;Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A;Title: Primary structure of the B-chain of human plasmin.
A;Reference number: A04627; MUID:77225245
A;Accession: A04627
A;Molecule type: protein
A;Residues: 581-810 <WI1>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla

A;Reference number: A04625; MUID:75093329
A;Accession: A04625
A;Molecule type: protein
A;Residues: 20-50,'Q',51-71,'E',73-85,87-100 <WI2>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A;Reference number: A04626; MUID:76043692
A;Accession: A04626
A;Molecule type: protein
A;Residues: 483-507,'E',509-604 <WI3>
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A;Title: The primary structure of human plasminogen. II. The histidine loop of human pla
A;Reference number: A92125; MUID:73149248
A;Contents: annotation; active site
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A;Reference number: A92048; MUID:69234739
A;Contents: annotation; active site
R;Trexler, M.; Vali, Z.; Patthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A;Reference number: A92382; MUID:82213905
A;Contents: annotation; omega-aminocarboxylic acid binding sites
R;Vali, Z.; Patthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A;Reference number: A92458; MUID:85054794
A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A;Title: Kringle domains of human angiostatin. Characterization of the anti-proliferativ
A;Reference number: A58811; MUID:97067211
A;Contents: annotation
R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A;Title: Generation of an angiostatin-like fragment from plasminogen by stromelysin-1 (M
A;Reference number: A58812; MUID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR
A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PMK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.
A;Reference number: A58819; MUID:92031502
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plas
A;Reference number: A58818; MUID:92031503
A;Contents: annotation
R;de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook,
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at
A;Reference number: A39483; MUID:92118803
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmi
A;Reference number: A58817; MUID:94237158
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH
d PIR:FGHUGB).
C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately
rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibit
C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial c
C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiostatin.
ting solid tumors.
C;Genetics:
A;Gene: GDB:PLG
A;Cross-references: GDB:119498; OMIM:173350
A;Map position: 6q26-6q27
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1;
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;20-96/Domain: activation peptide #status experimental <APT>

F;79-466/Product: angiostatin #status experimental <AST>
F;97-580,581-810/Product: plasmin #status experimental <MAT>
F;97-580/Domain: plasmin chain A #status experimental <CHA>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 100.0%; Score 34; DB 1; Length 810;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 549 RKLYDY 554

RESULT 2
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiostatin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of t
A;Reference number: A38514; MUID:91184812
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DEG>
A;Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LI2>
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many ot
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, both plasmin
mediately after dissociation from the clot. In the presence of the inhibitor, the activa
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiostatin. To
eful in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiostatin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 34; DB 1; Length 812;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 549 RKLYDY 554

RESULT 3
JQ1988
hypothetical 8.5K protein - turkey rhinotracheitis virus (strain UK/3BV/85)
N;Alternate names: ORF 2 protein
C;Species: turkey rhinotracheitis virus
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
C;Accession: JQ1988
R;Yu, Q.; Davis, P.J.; Brown, T.D.K.; Cavanagh, D.
J. Gen. Virol. 73, 1355-1363, 1992
A;Title: Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pn
A;Reference number: JQ1987; MUID:92300329
A;Accession: JQ1988
A;Molecule type: mRNA
A;Residues: 1-73 <YUQ>
A;Cross-references: GB:X63408; NID:g297846; PIDN:CAA45005.1; PID:g297848

Query Match 94.1%; Score 32; DB 2; Length 73;
Best Local Similarity 83.3%; Pred. No. 8.6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 25 RKIYDY 30

RESULT 4
B81293
hypothetical protein Cj1467 [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C;Accession: B81293
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A;Reference number: A81250; MUID:20150912
A;Accession: B81293
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <PAR>
A;Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73890.1; PID:g696
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1467

Query Match 94.1%; Score 32; DB 2; Length 254;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 210 RKIYDY 215

RESULT 5
S56296
probable membrane protein YFR041c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein F018
C;Species: Saccharomyces cerevisiae

C;Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 05-Dec-1997
C;Accession: S56296; S62252; S63788
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces ce
A;Reference number: S56186
A;Accession: S56296
A;Molecule type: DNA
A;Residues: 1-295 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PID:d1009921; PID:g836796; MIPS:YFR041c
R;Murakami, Y.
submitted to the EMBL Data Library, December 1994
A;Reference number: S62230
A;Accession: S62252
A;Molecule type: DNA
A;Residues: 1-295 <MUW>
A;Cross-references: EMBL:D44597; NID:g871938; PID:d1008597; PID:g871940
R;Eki, T.; Naitou, M.; Hagiwara, H.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiya,
Yeast 12, 149-167, 1996
A;Title: Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome V
A;Reference number: S63787; MUID:96287652
A;Accession: S63788
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-295 <EKI>
A;Cross-references: EMBL:D44597; NID:g871938; PID:d1008597; PID:g871940
C;Genetics:
A;Map position: 6R
A;Note: YFR041c
C;Superfamily: dnaJ amino-terminal homology
C;Keywords: transmembrane protein
F;8-24/Domain: transmembrane #status predicted <TM1>
F;44-108/Domain: dnaJ amino-terminal homology <DNJ>
F;133-149/Domain: transmembrane #status predicted <TM2>

Query Match 94.1%; Score 32; DB 2; Length 295;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 101 RKIYDY 106

RESULT 6
D82355
probable transposase VC0185 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: D82355
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: D82355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-404 <HEI>
A;Cross-references: GB:AE004108; GB:AE003852; NID:g9654578; PIDN:AAF93361.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0185
A;Map position: 1
C;Superfamily: Vibrio cholerae probable transposase VC0185

Query Match 94.1%; Score 32; DB 2; Length 404;
Best Local Similarity 83.3%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 273 RKMYDY 278

RESULT 7
A42289
glucose-fructose oxidoreductase (EC 1.1.-.-) precursor - Zymomonas mobilis
C;Species: Zymomonas mobilis
C;Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 08-Oct-1999
C;Accession: A42289
R;Kanagasundaram, V.; Scopes, R.K.
J. Bacteriol. 174, 1439-1447, 1992
A;Title: Cloning, sequence analysis, and expression of the structural gene encoding g
A;Reference number: A42289; MUID:92165717
A;Accession: A42289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <KAN>
A;Cross-references: GB:M97379; NID:g155587; PIDN:AAA27690.1; PID:g155588
C;Keywords: oxidoreductase

Query Match 94.1%; Score 32; DB 2; Length 439;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 133 RKIYDY 138

RESULT 8
A97409
hypothetical protein AGR_C_723 [imported] - Agrobacterium tumefaciens (strain C58, Ce
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C;Accession: A97409
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: A97409
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86226.1; PID:g15155328; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_723
A;Map position: circular chromosome

Query Match 91.2%; Score 31; DB 2; Length 142;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 129 RKVYDY 134

RESULT 9
E64358
ribosomal protein L5 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: E64358
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999
A;Accession: E64358
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-190 <BUL>
A;Cross-references: GB:U67497; GB:L77117; NID:g2826284; PIDN:AAB98458.1; PID:g1591171; T
C;Genetics:
A;Map position: FOR415069-415641
C;Superfamily: Escherichia coli ribosomal protein L5

Query Match 91.2%; Score 31; DB 2; Length 190;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
:|||||
Db 95 KKLYDY 100

RESULT 10
T29261
hypothetical protein C01G8.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29261
R;Du, Z.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A;Description: The sequence of C. elegans cosmid C01G8.
A;Reference number: Z20597
A;Accession: T29261
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-345 <DUZ>
A;Cross-references: EMBL:U80439; PIDN:AAB37641.1; GSPDB:GN00019; CESP:C01G8.4
A;Experimental source: strain Bristol N2; clone C01G8
C;Genetics:
A;Gene: CESP:C01G8.4
A;Map position: 1
A;Introns: 18/2; 54/3; 80/2; 119/2; 169/2; 207/3

Query Match 91.2%; Score 31; DB 2; Length 345;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
|:||||
Db 85 RRLYDY 90

RESULT 11
S26824
hypothetical protein - phage Cf1c
C;Species: phage Cf1c
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C;Accession: S26824
R;Kuo, T.T.; Tan, M.S.; Su, M.T.; Yang, M.K.
Nucleic Acids Res. 19, 2498, 1991
A;Title: Complete nucleotide sequence of filamentous phage Cf1c from Xanthomonas campest
A;Reference number: S26824; MUID:91252299
A;Accession: S26824
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-388 <KUO>
A;Cross-references: EMBL:M57538; NID:g166169; PIDN:AAA32201.1; PID:g166170
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1991

Query Match 91.2%; Score 31; DB 2; Length 388;
Best Local Similarity 83.3%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
:|||||
Db 155 KKLYDY 160

RESULT 12
B89954
hypothetical protein SA1524 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: B89954
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: B89954
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-409 <KUR>
A;Cross-references: GB:BA000018; PID:g13701497; PIDN:BAB42791.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA1524
C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 91.2%; Score 31; DB 2; Length 409;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||:|||
Db 52 RKVYDY 57

RESULT 13
E70207
antigen S1 - Lyme disease spirochete plasmid A/lp54
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: E70207; I40296
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: E70207
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-417 <KLE>
A;Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66229.1; PID:g2690228; TIGR:BB
A;Experimental source: strain B31
R;Feng, S.; Das, S.; Lam, T.; Flavell, R.A.; Fikrig, E.
Infect. Immun. 63, 3459-3466, 1995
A;Title: A 55-kilodalton antigen encoded by a gene on a Borrelia burgdorferi 49-kilob
A;Reference number: I40296; MUID:95369900
A;Accession: I40296
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18,'R',20-417 <RES>
A;Cross-references: GB:L34016; NID:g1063416; PIDN:AAA81351.1; PID:g1063417
A;Experimental source: strain N40
C;Genetics:
A;Gene: S1
A;Genome: plasmid

Query Match 91.2%; Score 31; DB 2; Length 417;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
:|||||
Db 397 KKLYDY 402

RESULT 14

E95130
enolase [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: E95130
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: E95130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75238.1; PID:g14972605; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1128
C;Superfamily: enolase

  Query Match 91.2%; Score 31; DB 2; Length 434;
  Best Local Similarity 83.3%; Pred. No. 85;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
  ||:|||
Db 253 RKVYDY 258

RESULT 15
D98001
phosphopyruvate hydratase (EC 4.2.1.11) [imported] - Streptococcus pneumoniae (strain R6
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: D98001
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D98001
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99840.1; PID:g15458655; GSPDB:GN00174
C;Genetics:
A;Gene: eno
C;Superfamily: enolase
C;Keywords: carbon-oxygen lyase; hydro-lyase

  Query Match 91.2%; Score 31; DB 2; Length 434;
  Best Local Similarity 83.3%; Pred. No. 85;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
  ||:|||
Db 253 RKVYDY 258

Search completed: November 8, 2002, 09:36:45
Job time : 10 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:34:41 ; Search time 30 Seconds
(without alignments)
22.215 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 6
Sequence: 1 RKLYDY 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	100.0	6	18	AAW34302	Kringle 5 peptide
2	6	100.0	6	21	AAB01903	Human plasminogen
3	6	100.0	6	22	AAB92098	Laminin fragment S
4	6	100.0	6	22	AAB36572	Mammalian kringle
5	6	100.0	7	18	AAW34303	Kringle 5 peptide
6	6	100.0	7	18	AAW34304	Kringle 5 peptide
7	6	100.0	7	18	AAW34306	Kringle 5 peptide
8	6	100.0	7	18	AAW34299	Kringle 5 peptide
9	6	100.0	7	18	AAW34290	Human kringle 5 pe
10	6	100.0	7	18	AAW34291	Human kringle 5 pe
11	6	100.0	7	21	AAB01888	Human plasminogen
12	6	100.0	7	21	AAB01889	Human plasminogen
13	6	100.0	7	21	AAB01900	Human plasminogen
14	6	100.0	7	22	AAB92097	Laminin fragment S
15	6	100.0	7	22	AAB36571	Mammalian kringle
16	6	100.0	7	22	AAB36576	Mammalian kringle
17	6	100.0	8	22	AAB83292	Kringle peptide #1
18	6	100.0	8	22	AAB92092	Laminin fragment S
19	6	100.0	8	22	AAB92101	Laminin fragment S
20	6	100.0	8	22	AAB36567	Mammalian kringle
21	6	100.0	8	22	AAB36575	Mammalian kringle
22	6	100.0	11	18	AAW34298	Kringle 5 peptide
23	6	100.0	11	21	AAB01899	Human plasminogen
24	6	100.0	11	22	AAB92094	Laminin fragment S
25	6	100.0	11	22	AAB36569	Mammalian kringle
26	6	100.0	12	22	AAB92089	Laminin fragment S
27	6	100.0	12	22	AAB92093	Laminin fragment S
28	6	100.0	12	22	AAB36564	Mammalian kringle
29	6	100.0	12	22	AAB36568	Mammalian kringle
30	6	100.0	21	21	AAB15521	Sequences flanking
31	6	100.0	23	22	AAB83293	Kringle peptide #2
32	6	100.0	23	22	AAB92096	Laminin fragment S
33	6	100.0	23	22	AAB36570	Mammalian kringle
34	6	100.0	24	22	AAB92090	Laminin fragment S
35	6	100.0	24	22	AAB92095	Laminin fragment S
36	6	100.0	24	22	AAB36565	Mammalian kringle
37	6	100.0	79	18	AAW19256	Human plasminogen
38	6	100.0	85	21	AAB08413	Amino acid sequenc
39	6	100.0	90	21	AAB01914	Human plasminogen
40	6	100.0	91	21	AAY58868	Human plasminogen
41	6	100.0	93	21	AAB01917	Human plasminogen
42	6	100.0	95	21	AAB01913	Human plasminogen
43	6	100.0	98	21	AAB01916	Human plasminogen
44	6	100.0	101	18	AAW34286	Human kringle 5 pe
45	6	100.0	101	21	AAB01890	Human plasminogen

ALIGNMENTS

RESULT 1
AAW34302
ID AAW34302 standard; peptide; 6 AA.
XX
AC AAW34302;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Arg"
FT Modified-site 6
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.
XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.

XX
PI Davidson DJ, Gubbins EJ, Wang J;
XX
DR WPI; 1997-558670/51.
XX
PT New kringle 5 peptide(s) and fusion proteins derived from
PT plasminogen - useful as anti-angiogenesis agents for treating
PT cancer, psoriasis, arthritis etc., including gene therapy
XX
PS Example 9; Page 45; 78pp; English.
XX
CC This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC plasminogen. K5 peptide fragments homologous to this sequence, are
CC anti-angiogenesis agents, specifically for treating or preventing cancer,
CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC lymphomas, haemangiomas. They can also be used for treating or preventing
CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC fragments can also be used to treat autoimmune or ocular diseases,
CC capillary proliferation within atherosclerotic plaque, haemophiliac
CC joints, wound granulation, ulcers etc., also as contraceptives that
CC inhibit ovulation and establishment of the placenta. K5 antisera or
CC (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC for diagnosis and for affinity purification of K5 receptors. The K5
CC receptors may then be expressed in tumour cells to increase their
CC response to the peptides or used for identification of smaller
CC antagonists. The Ab are used to detect/quantify the peptides in
CC biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC inhibit proliferation of endothelial cells with low toxicity against
CC normal cells. Typically they have 800-times greater inhibitory activity
CC against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 6; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 2
AAB01903
ID AAB01903 standard; peptide; 6 AA.
XX
AC AAB01903;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5 peptide fragment #9.
XX
KW Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl moiety"
FT Modified-site 6
FT /note= "C-terminal amide moiety"
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 9; Column 37; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC Syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC plasminogen kringle 5-derived peptides synthesised and used in
CC exemplifications of the invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 6; DB 21; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 3
AAB92098
ID AAB92098 standard; Peptide; 6 AA.
XX
AC AAB92098;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1274.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 612; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 6; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 4
AAB36572
ID AAB36572 standard; Peptide; 6 AA.
XX
AC AAB36572;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:11.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 6; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 6; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 5
AAW34303
ID AAW34303 standard; peptide; 7 AA.
XX
AC AAW34303;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Pro"
FT Modified-site 5
FT /note= "3-I-tyrosyl"
FT Modified-site 7
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.

```
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 10; Page 45; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 6
AAW34304
ID   AAW34304 standard; peptide; 7 AA.
XX
AC   AAW34304;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-Ac-Pro"
FT   Modified-site   7
FT                   /note= "3-I-tyrosyl"
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 11; Page 45; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 7
AAW34306
ID   AAW34306 standard; peptide; 7 AA.
XX
AC   AAW34306;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
```

```
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-Ac-Pro"
FT   Modified-site   7
FT                   /note= "3-I-tyrosyl"
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 13; Page 46; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 8
AAW34299
ID   AAW34299 standard; peptide; 7 AA.
XX
AC   AAW34299;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-Ac-Pro"
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 6; Page 44; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 9
```

```
AAW34290
ID   AAW34290 standard; peptide; 7 AA.
XX
AC   AAW34290;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "optionally has nitrogen protecting group"
FT   Modified-site   7
FT                   /note= "3-I-Tyrosyl"
FT   Misc-difference 7
FT                   /note= "optionally has carboxylic acid protecting group"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Claim 62; Page 67; 78pp; English.
XX
CC   This sequence is a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 10
AAW34291
ID   AAW34291 standard; peptide; 7 AA.
XX
AC   AAW34291;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "optionally has nitrogen protecting group"
FT   Modified-site   5
FT                   /note= "3-I-Tyrosyl"
FT   Misc-difference 7
FT                   /note= "optionally has carboxylic acid protecting group"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Claim 62; Page 67; 78pp; English.
XX
CC   This sequence is a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
```

```
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels  .  0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 11
AAB01888
ID   AAB01888 standard; peptide; 7 AA.
XX
AC   AAB01888;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5-derived peptide, SEQ ID NO:6.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   5
FT                   /note= "3-Iodotyrosine"
FT   Modified-site   7
FT                   /note= "C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 10; Column 38; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 12
AAB01889
ID   AAB01889 standard; peptide; 7 AA.
XX
AC   AAB01889;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5-derived peptide, SEQ ID NO:18.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   7
FT                   /note= "3-Iodotyrosine; C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 11; Column 38; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
```

```
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 13
AAB01900
ID   AAB01900 standard; peptide; 7 AA.
XX
AC   AAB01900;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 peptide fragment #6.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   7
FT                   /note= "C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 6; Column 36; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 6;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 14
AAB92097
ID   AAB92097 standard; Peptide; 7 AA.
XX
AC   AAB92097;
XX
DT   22-JUN-2001  (first entry)
XX
DE   Laminin fragment SEQ ID NO:1273.
XX
KW   Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW   blood component; modification; succinimidyl; maleimido group; amino;
KW   hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200069900-A2.
XX
PD   23-NOV-2000.
XX
PF   17-MAY-2000; 2000WO-US13576.
XX
PR   17-MAY-1999;   99US-0134406.
PR   10-SEP-1999;   99US-0153406.
PR   15-OCT-1999;   99US-0159783.
XX
PA   (CONJ-) CONJUCHEM INC.
XX
PI   Bridon DP,  Ezrin AM,  Milner PG,  Holmes DL,  Thibaudeau K;
XX
DR   WPI; 2001-112059/12.
XX
PT   Modifying and attaching therapeutic peptides to albumin prevents
PT   peptidase degradation, useful for increasing length of in vivo activity
PT   -
XX
PS   Disclosure; Page 612; 733pp; English.
XX
CC   The present invention describes a modified therapeutic peptide (I)
CC   comprising a therapeutically active amino acid region (III) and a
CC   reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC   a less therapeutically active amino acid region (IV), which covalently
CC   bonds with amino/hydroxyl/thiol groups on blood components to form a
```

CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 6; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 15
AAB36571
ID AAB36571 standard; Peptide; 7 AA.
XX
AC AAB36571;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:10.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 6; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 6; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

Search completed: November 8, 2002, 09:37:20
Job time : 30 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:21 ; Search time 17.6667 Seconds
(without alignments)
37.723 Million cell updates/sec

Title: US-09-657-431-11
Perfect score: 34
Sequence: 1 RKLYDY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	6	18	AAW34302	Kringle 5 peptide
2	34	100.0	6	21	AAB01903	Human plasminogen
3	34	100.0	6	22	AAB92098	Laminin fragment S
4	34	100.0	6	22	AAB36572	Mammalian kringle
5	34	100.0	7	18	AAW34303	Kringle 5 peptide
6	34	100.0	7	18	AAW34304	Kringle 5 peptide
7	34	100.0	7	18	AAW34306	Kringle 5 peptide
8	34	100.0	7	18	AAW34299	Kringle 5 peptide
9	34	100.0	7	18	AAW34290	Human kringle 5 pe
10	34	100.0	7	18	AAW34291	Human kringle 5 pe
11	34	100.0	7	21	AAB01888	Human plasminogen
12	34	100.0	7	21	AAB01889	Human plasminogen
13	34	100.0	7	21	AAB01900	Human plasminogen
14	34	100.0	7	22	AAB92097	Laminin fragment S
15	34	100.0	7	22	AAB36571	Mammalian kringle
16	34	100.0	7	22	AAB36576	Mammalian kringle
17	34	100.0	8	22	AAB83292	Kringle peptide #1
18	34	100.0	8	22	AAB92092	Laminin fragment S
19	34	100.0	8	22	AAB92101	Laminin fragment S
20	34	100.0	8	22	AAB36567	Mammalian kringle
21	34	100.0	8	22	AAB36575	Mammalian kringle
22	34	100.0	11	18	AAW34298	Kringle 5 peptide
23	34	100.0	11	21	AAB01899	Human plasminogen
24	34	100.0	11	22	AAB92094	Laminin fragment S
25	34	100.0	11	22	AAB36569	Mammalian kringle
26	34	100.0	12	22	AAB92089	Laminin fragment S
27	34	100.0	12	22	AAB92093	Laminin fragment S
28	34	100.0	12	22	AAB36564	Mammalian kringle
29	34	100.0	12	22	AAB36568	Mammalian kringle
30	34	100.0	21	21	AAB15521	Sequences flanking
31	34	100.0	23	22	AAB83293	Kringle peptide #2
32	34	100.0	23	22	AAB92096	Laminin fragment S
33	34	100.0	23	22	AAB36570	Mammalian kringle
34	34	100.0	24	22	AAB92090	Laminin fragment S
35	34	100.0	24	22	AAB92095	Laminin fragment S
36	34	100.0	24	22	AAB36565	Mammalian kringle
37	34	100.0	79	18	AAW19256	Human plasminogen
38	34	100.0	85	21	AAB08413	Amino acid sequenc
39	34	100.0	90	21	AAB01914	Human plasminogen
40	34	100.0	91	21	AAY58868	Human plasminogen
41	34	100.0	93	21	AAB01917	Human plasminogen
42	34	100.0	95	21	AAB01913	Human plasminogen
43	34	100.0	98	21	AAB01916	Human plasminogen
44	34	100.0	101	18	AAW34286	Human kringle 5 pe
45	34	100.0	101	21	AAB01890	Human plasminogen

ALIGNMENTS

RESULT 1
AAW34302
ID AAW34302 standard; peptide; 6 AA.
XX
AC AAW34302;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Arg"
FT Modified-site 6
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.
XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.

XX
PI Davidson DJ, Gubbins EJ, Wang J;
XX
DR WPI; 1997-558670/51.
XX
PT New kringle 5 peptide(s) and fusion proteins derived from
PT plasminogen - useful as anti-angiogenesis agents for treating
PT cancer, psoriasis, arthritis etc., including gene therapy
XX
PS Example 9; Page 45; 78pp; English.
XX
CC This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC plasminogen. K5 peptide fragments homologous to this sequence, are
CC anti-angiogenesis agents, specifically for treating or preventing cancer,
CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC lymphomas, haemangiomas. They can also be used for treating or preventing
CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC fragments can also be used to treat autoimmune or ocular diseases,
CC capillary proliferation within atherosclerotic plaque, haemophiliac
CC joints, wound granulation, ulcers etc., also as contraceptives that
CC inhibit ovulation and establishment of the placenta. K5 antisera or
CC (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC for diagnosis and for affinity purification of K5 receptors. The K5
CC receptors may then be expressed in tumour cells to increase their
CC response to the peptides or used for identification of smaller
CC antagonists. The Ab are used to detect/quantify the peptides in
CC biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC inhibit proliferation of endothelial cells with low toxicity against
CC normal cells. Typically they have 800-times greater inhibitory activity
CC against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 2
AAB01903
ID AAB01903 standard; peptide; 6 AA.
XX
AC AAB01903;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5 peptide fragment #9.
XX
KW Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl moiety"
FT Modified-site 6
FT /note= "C-terminal amide moiety"
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 9; Column 37; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC Syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC plasminogen kringle 5-derived peptides synthesised and used in
CC exemplifications of the invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 21; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 3
AAB92098
ID AAB92098 standard; Peptide; 6 AA.
XX
AC AAB92098;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1274.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 612; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 4
AAB36572
ID AAB36572 standard; Peptide; 6 AA.
XX
AC AAB36572;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:11.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 6; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 5
AAW34303
ID AAW34303 standard; peptide; 7 AA.
XX
AC AAW34303;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Pro"
FT Modified-site 5
FT /note= "3-I-tyrosyl"
FT Modified-site 7
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.

XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ, Gubbins EJ, Wang J;
XX
DR WPI; 1997-558670/51.
XX
PT New kringle 5 peptide(s) and fusion proteins derived from
PT plasminogen - useful as anti-angiogenesis agents for treating
PT cancer, psoriasis, arthritis etc., including gene therapy
XX
PS Example 10; Page 45; 78pp; English.
XX
CC This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC plasminogen. K5 peptide fragments homologous to this sequence, are
CC anti-angiogenesis agents, specifically for treating or preventing cancer,
CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC lymphomas, haemangiomas. They can also be used for treating or preventing
CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC fragments can also be used to treat autoimmune or ocular diseases,
CC capillary proliferation within atherosclerotic plaque, haemophiliac
CC joints, wound granulation, ulcers etc., also as contraceptives that
CC inhibit ovulation and establishment of the placenta. K5 antisera or
CC (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC for diagnosis and for affinity purification of K5 receptors. The K5
CC receptors may then be expressed in tumour cells to increase their
CC response to the peptides or used for identification of smaller
CC antagonists. The Ab are used to detect/quantify the peptides in
CC biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC inhibit proliferation of endothelial cells with low toxicity against
CC normal cells. Typically they have 800-times greater inhibitory activity
CC against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 34; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 2 RKLYDY 7

RESULT 6
AAW34304
ID AAW34304 standard; peptide; 7 AA.
XX
AC AAW34304;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-Ac-Pro"
FT Modified-site 7
FT /note= "3-I-tyrosyl"
FT Modified-site 7
FT /note= "C-terminal amide"
XX
PN WO9741824-A2.
XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ, Gubbins EJ, Wang J;
XX
DR WPI; 1997-558670/51.
XX
PT New kringle 5 peptide(s) and fusion proteins derived from
PT plasminogen - useful as anti-angiogenesis agents for treating
PT cancer, psoriasis, arthritis etc., including gene therapy
XX
PS Example 11; Page 45; 78pp; English.
XX
CC This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC plasminogen. K5 peptide fragments homologous to this sequence, are
CC anti-angiogenesis agents, specifically for treating or preventing cancer,
CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC lymphomas, haemangiomas. They can also be used for treating or preventing
CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC fragments can also be used to treat autoimmune or ocular diseases,
CC capillary proliferation within atherosclerotic plaque, haemophiliac
CC joints, wound granulation, ulcers etc., also as contraceptives that
CC inhibit ovulation and establishment of the placenta. K5 antisera or
CC (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC for diagnosis and for affinity purification of K5 receptors. The K5
CC receptors may then be expressed in tumour cells to increase their
CC response to the peptides or used for identification of smaller
CC antagonists. The Ab are used to detect/quantify the peptides in
CC biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC inhibit proliferation of endothelial cells with low toxicity against
CC normal cells. Typically they have 800-times greater inhibitory activity
CC against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 34; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 2 RKLYDY 7

RESULT 7
AAW34306
ID AAW34306 standard; peptide; 7 AA.
XX
AC AAW34306;
XX
DT 14-MAY-1998 (first entry)
XX
DE Kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;

```
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-Ac-Pro"
FT   Modified-site   7
FT                   /note= "3-I-tyrosyl"
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 13; Page 46; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 8
AAW34299
ID   AAW34299 standard; peptide; 7 AA.
XX
AC   AAW34299;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-Ac-Pro"
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Example 6; Page 44; 78pp; English.
XX
CC   This sequence is synthetic a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 9
```

```
AAW34290
ID   AAW34290 standard; peptide; 7 AA.
XX
AC   AAW34290;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "optionally has nitrogen protecting group"
FT   Modified-site   7
FT                   /note= "3-I-Tyrosyl"
FT   Misc-difference 7
FT                   /note= "optionally has carboxylic acid protecting group"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Claim 62; Page 67; 78pp; English.
XX
CC   This sequence is a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RKLYDY 6
           ||||||
Db       2 RKLYDY 7


RESULT 10
AAW34291
ID   AAW34291 standard; peptide; 7 AA.
XX
AC   AAW34291;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human kringle 5 peptide fragment.
XX
KW   Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW   metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW   psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW   autoimmune disease; ocular disease; capillary proliferation; therapy;
KW   kringle 5 receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "optionally has nitrogen protecting group"
FT   Modified-site   5
FT                   /note= "3-I-Tyrosyl"
FT   Misc-difference 7
FT                   /note= "optionally has carboxylic acid protecting group"
XX
PN   WO9741824-A2.
XX
PD   13-NOV-1997.
XX
PF   05-MAY-1997;   97WO-US07700.
XX
PR   03-APR-1997;   97US-0832087.
PR   03-MAY-1996;   96US-0643219.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ,  Gubbins EJ,  Wang J;
XX
DR   WPI; 1997-558670/51.
XX
PT   New kringle 5 peptide(s) and fusion proteins derived from
PT   plasminogen - useful as anti-angiogenesis agents for treating
PT   cancer, psoriasis, arthritis etc., including gene therapy
XX
PS   Claim 62; Page 67; 78pp; English.
XX
CC   This sequence is a kringle 5 (K5) peptide homologous to human
CC   plasminogen. K5 peptide fragments homologous to this sequence, are
CC   anti-angiogenesis agents, specifically for treating or preventing cancer,
CC   particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC   lymphomas, haemangiomas. They can also be used for treating or preventing
CC   psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC   fragments can also be used to treat autoimmune or ocular diseases,
CC   capillary proliferation within atherosclerotic plaque, haemophiliac
CC   joints, wound granulation, ulcers etc., also as contraceptives that
CC   inhibit ovulation and establishment of the placenta. K5 antisera or
CC   (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC   agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC   healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC   for diagnosis and for affinity purification of K5 receptors. The K5
CC   receptors may then be expressed in tumour cells to increase their
CC   response to the peptides or used for identification of smaller
CC   antagonists. The Ab are used to detect/quantify the peptides in
CC   biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC   inhibit proliferation of endothelial cells with low toxicity against
CC   normal cells. Typically they have 800-times greater inhibitory activity
```

```
CC   against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 11
AAB01888
ID   AAB01888 standard; peptide; 7 AA.
XX
AC   AAB01888;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5-derived peptide, SEQ ID NO:6.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   5
FT                   /note= "3-Iodotyrosine"
FT   Modified-site   7
FT                   /note= "C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 10; Column 38; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 12
AAB01889
ID   AAB01889 standard; peptide; 7 AA.
XX
AC   AAB01889;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5-derived peptide, SEQ ID NO:18.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   7
FT                   /note= "3-Iodotyrosine; C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 11; Column 38; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
```

```
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 13
AAB01900
ID   AAB01900 standard; peptide; 7 AA.
XX
AC   AAB01900;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 peptide fragment #6.
XX
KW   Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl moiety"
FT   Modified-site   7
FT                   /note= "C-terminal amide moiety"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 6; Column 36; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01888, AAB01889 and AAB01895-B01905 represent human
CC   plasminogen kringle 5-derived peptides synthesised and used in
CC   exemplifications of the invention.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 34;  DB 21;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RKLYDY 6
          ||||||
Db      2 RKLYDY 7


RESULT 14
AAB92097
ID   AAB92097 standard; Peptide; 7 AA.
XX
AC   AAB92097;
XX
DT   22-JUN-2001  (first entry)
XX
DE   Laminin fragment SEQ ID NO:1273.
XX
KW   Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW   blood component; modification; succinimidyl; maleimido group; amino;
KW   hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200069900-A2.
XX
PD   23-NOV-2000.
XX
PF   17-MAY-2000; 2000WO-US13576.
XX
PR   17-MAY-1999;   99US-0134406.
PR   10-SEP-1999;   99US-0153406.
PR   15-OCT-1999;   99US-0159783.
XX
PA   (CONJ-) CONJUCHEM INC.
XX
PI   Bridon DP,  Ezrin AM,  Milner PG,  Holmes DL,  Thibaudeau K;
XX
DR   WPI; 2001-112059/12.
XX
PT   Modifying and attaching therapeutic peptides to albumin prevents
PT   peptidase degradation, useful for increasing length of in vivo activity
PT   -
XX
PS   Disclosure; Page 612; 733pp; English.
XX
CC   The present invention describes a modified therapeutic peptide (I)
CC   comprising a therapeutically active amino acid region (III) and a
CC   reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC   a less therapeutically active amino acid region (IV), which covalently
CC   bonds with amino/hydroxyl/thiol groups on blood components to form a
```

CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 34; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

RESULT 15
AAB36571
ID AAB36571 standard; Peptide; 7 AA.
XX
AC AAB36571;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:10.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 6; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 34; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKLYDY 6
||||||
Db 1 RKLYDY 6

Search completed: November 8, 2002, 09:34:52
Job time : 18.6667 secs

(5)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:22 ; Search time 28.6667 Seconds
(without alignments)
72.416 Million cell updates/sec

Title: US-09-657-431-5
Perfect score: 75
Sequence: 1 RNPDGDVGGPWK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	70	93.3	334	6	O46507	O46507 papio hamad
2	70	93.3	810	4	Q15146	Q15146 homo sapien
3	64	85.3	567	4	Q13208	Q13208 homo sapien
4	64	85.3	648	4	Q9H1V4	Q9h1v4 homo sapien
5	64	85.3	812	11	Q9R0W3	Q9r0w3 rattus norv
6	64	85.3	812	11	Q91WJ5	Q91wj5 mus musculu
7	58	77.3	806	6	O18783	O18783 macropus eu
8	56	74.7	716	11	P70521	P70521 rattus norv
9	56	74.7	716	11	Q91XG8	Q91xg8 mus musculu
10	55	73.3	709	13	Q90ZN6	Q90zn6 brachydanio
11	53	70.7	710	13	Q91402	Q91402 xenopus. he
12	52	69.3	75	6	Q9BGN9	Q9bgn9 bos taurus
13	52	69.3	109	6	Q9N1B8	Q9n1b8 ovis aries
14	52	69.3	208	4	Q9BYM0	Q9bym0 homo sapien
15	52	69.3	210	4	Q13494	Q13494 homo sapien
16	52	69.3	211	11	O55027	O55027 mus musculu
17	52	69.3	290	4	Q02935	Q02935 homo sapien
18	52	69.3	296	4	Q14519	Q14519 homo sapien
19	52	69.3	726	13	Q90978	Q90978 gallus gall
20	52	69.3	728	6	Q9BH09	Q9bh09 felis silve
21	50	66.7	111	6	O77688	O77688 oryctolagus
22	50	66.7	313	13	Q9PU78	Q9pu78 crocodylus
23	50	66.7	704	13	Q90865	Q90865 gallus gall
24	50	66.7	716	13	Q91691	Q91691 xenopus lae
25	49	65.3	215	13	O42341	O42341 gallus gall
26	48	64.0	263	4	O00318	O00318 homo sapien
27	48	64.0	263	4	Q96FE7	Q96fe7 homo sapien
28	48	64.0	717	13	P70006	P70006 xenopus lae
29	47.5	63.3	117	4	Q9UGS5	Q9ugs5 homo sapien
30	47.5	63.3	452	13	Q90Y90	Q90y90 xenopus lae
31	47.5	63.3	473	4	Q9BY70	Q9by70 homo sapien
32	47.5	63.3	473	11	Q99N43	Q99n43 mus musculu
33	47.5	63.3	473	11	Q924S4	Q924s4 rattus norv
34	47.5	63.3	492	4	Q96MU8	Q96mu8 homo sapien
35	47	62.7	382	13	Q90WT4	Q90wt4 crocodylus
36	47	62.7	385	5	Q25101	Q25101 herdmania m
37	47	62.7	420	13	Q90504	Q90504 eptatretus
38	46	61.3	260	16	Q57371	Q57371 borrelia bu
39	46	61.3	312	4	Q9NSV1	Q9nsv1 homo sapien
40	45.5	60.7	399	4	Q96GL8	Q96gl8 homo sapien
41	45.5	60.7	420	4	Q9BTP9	Q9btp9 homo sapien
42	45	60.0	140	11	Q9D874	Q9d874 mus musculu
43	45	60.0	607	13	Q91001	Q91001 gallus gall
44	45	60.0	608	13	Q9PTW7	Q9ptw7 struthio ca
45	44	58.7	522	10	Q9LIW0	Q9liw0 oryza sativ

ALIGNMENTS

```
RESULT 1
O46507
ID   O46507      PRELIMINARY;      PRT;   334 AA.
AC   O46507;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN (FRAGMENT).
GN   BABPEPSG.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Cox L.A., Jett C., Hixson J.E.;
RT   "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT   Site Mutation is Associated with Deletion of a Single Exon in a Null
RT   Allele.";
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF029692; AAB97887.1; -.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
```

```
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  36791 MW;  C7DC06E03B965286 CRC64;

  Query Match             93.3%;  Score 70;  DB 6;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 0.0019;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db     56 RNPDGDVGGPW 66


RESULT 2
Q15146
ID   Q15146      PRELIMINARY;      PRT;   810 AA.
AC   Q15146;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA   Mitchell D., Robinson J.H.;
RT   "Expression of recombinant human plasminogen and aglycoplasminogen in
RT   HeLa cells.";
RL   Fibrinolysis 0:0-0(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; M74220; AAA36451.1; -.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    810       PLASMINOGEN.
SQ   SEQUENCE   810 AA;  90555 MW;  B05C7D4B0D020B3C CRC64;

  Query Match             93.3%;  Score 70;  DB 4;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 0.0048;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db    532 RNPDGDVGGPW 542


RESULT 3
Q13208
ID   Q13208      PRELIMINARY;      PRT;   567 AA.
AC   Q13208;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN HOMOLOG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20191171; PubMed=10728827;
RA   Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
RA   Carritt B.;
RT   "Structure of the human D1F15S1A locus: a chromosome 1 locus with 97%
RT   identity to the chromosome 3 gene coding for hepatocyte growth factor-
RT   like protein.";
RL   DNA Seq. 8:409-413(1998).
DR   EMBL; U28054; AAC63092.1; -.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.977; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   567 AA;  64117 MW;  3FC38B07F1645810 CRC64;

  Query Match             85.3%;  Score 64;  DB 4;  Length 567;
  Best Local Similarity   90.9%;  Pred. No. 0.031;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| |||| 
Db    133 RNPDGDPGGPW 143


RESULT 4
Q9H1V4
ID   Q9H1V4      PRELIMINARY;      PRT;   648 AA.
AC   Q9H1V4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DJ1182A14.3 (SIMILAR TO MST1 (MACROPHAGE STIMULATING 1 (HEPATOCYTE
DE   GROWTH FACTOR-LIKE))).
GN   DJ1182A14.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bird C.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AL137798; CAC17639.1; -.
DR   HSSP; P00747; 5HPG.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
```

```
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   648 AA;  72781 MW;  4CE077057350E463 CRC64;

  Query Match             85.3%;  Score 64;  DB 4;  Length 648;
  Best Local Similarity   90.9%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           |||||| ||||
Db     113 RNPDGDPGGPW 123


RESULT 5
Q9R0W3
ID   Q9R0W3      PRELIMINARY;      PRT;   812 AA.
AC   Q9R0W3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PROTEIN PRECURSOR (EC 3.4.21.7).
GN   PLASMINOGEN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Bangert K., Johnsen A.H., Thorsen S.;
RT   "Rat plasminogen: cDNA and gene structure.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=91250378; PubMed=1645711;
RA   Kanalas J.J., Makker S.P.;
RT   "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT   receptor site for plasminogen.";
RL   J. Biol. Chem. 266:10825-10829(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AJ242649; CAB46014.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
SQ   SEQUENCE   812 AA;  90535 MW;  8C703C51410EBC9E CRC64;

  Query Match             85.3%;  Score 64;  DB 11;  Length 812;
  Best Local Similarity   90.9%;  Pred. No. 0.045;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||| |||
Db     532 RNPDGDVNGPW 542


RESULT 6
Q91WJ5
ID   Q91WJ5      PRELIMINARY;      PRT;   812 AA.
AC   Q91WJ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014773; AAH14773.1; -.
SQ   SEQUENCE   812 AA;  90781 MW;  24173260E6A2FFD2 CRC64;

  Query Match             85.3%;  Score 64;  DB 11;  Length 812;
  Best Local Similarity   90.9%;  Pred. No. 0.045;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||| |||
Db     532 RNPDGDVNGPW 542


RESULT 7
O18783
ID   O18783      PRELIMINARY;      PRT;   806 AA.
AC   O18783;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Macropus eugenii (Tammar wallaby).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=98004511; PubMed=9342350;
RA   Lawn R.M., Schwartz K., Patthy L.;
RT   "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF012297; AAB65760.1; -.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
```

```
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   806 AA;  90981 MW;  95FAA86DC20064D5 CRC64;

  Query Match             77.3%;  Score 58;  DB 6;  Length 806;
  Best Local Similarity   81.8%;  Pred. No. 0.42;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||  |||
Db    527 RNPDGDPNGPW 537


RESULT 8
P70521
ID   P70521      PRELIMINARY;      PRT;   716 AA.
AC   P70521;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MACROPHAGE STIMULATING PROTEIN PRECURSOR.
GN   MSP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=97011126; PubMed=8858136;
RA   Ohshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA   Takasu N., Suda T.;
RT   "Molecular cloning of Rat Macrophage-stimulating protein and its
RT   involvement in the Male Reproductive System.";
RL   Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; X95096; CAA64473.1; -.
DR   HSSP; P00747; 1KRN.
DR   MEROPS; S01.975; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     31       POTENTIAL.
SQ   SEQUENCE   716 AA;  80733 MW;  06B7DF3EF56D921F CRC64;

  Query Match             74.7%;  Score 56;  DB 11;  Length 716;
  Best Local Similarity   81.8%;  Pred. No. 0.78;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||  |||
Db    429 RNPDGDSHGPW 439


RESULT 9
Q91XG8
ID   Q91XG8      PRELIMINARY;      PRT;   716 AA.
AC   Q91XG8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010551; AAH10551.1; -.
SQ   SEQUENCE   716 AA;  80693 MW;  12474C48A7D4B46D CRC64;

  Query Match             74.7%;  Score 56;  DB 11;  Length 716;
  Best Local Similarity   81.8%;  Pred. No. 0.78;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||  |||
Db    429 RNPDGDSHGPW 439


RESULT 10
Q90ZN6
ID   Q90ZN6      PRELIMINARY;      PRT;   709 AA.
AC   Q90ZN6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE 1.
GN   HGFL1.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bassett D.I., Wilson S.W.;
RT   "Early expression of zebrafish Hepatocyte Growth Factor-Like 1
RT   suggests a conserved role in vertebrate neural induction.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF370035; AAK54207.1; -.
SQ   SEQUENCE   709 AA;  81271 MW;  9907236C5DB73A20 CRC64;

  Query Match             73.3%;  Score 55;  DB 13;  Length 709;
  Best Local Similarity   81.8%;  Pred. No. 1.1;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||  |||
Db    422 RNPDGDHHGPW 432


RESULT 11
Q91402
```

```
ID   Q91402      PRELIMINARY;      PRT;   710 AA.
AC   Q91402;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR.
GN   HGF.
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
OX   NCBI_TaxID=8353;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TAILBUD;
RX   MEDLINE=95267690; PubMed=7748783;
RA   Nakamura H., Tashiro K., Nakamura T., Shiokawa K.;
RT   "Molecular cloning of Xenopus HGF cDNA and its expression studies in
RT   Xenopus early embryogenesis.";
RL   Mech. Dev. 49:123-131(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; S77422; AAB34354.2; -.
DR   HSSP; P14210; 1BHT.
DR   MEROPS; S01.976; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 3.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   710 AA;  81487 MW;  5FE6480BE31C27FC CRC64;

  Query Match             70.7%;  Score 53;  DB 13;  Length 710;
  Best Local Similarity   72.7%;  Pred. No. 2.4;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     165 RNPKGEEGGPW 175


RESULT 12
Q9BGN9
ID   Q9BGN9      PRELIMINARY;      PRT;    75 AA.
AC   Q9BGN9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RA   Murakami S., Fujiwara C., Miyamoto Y., Takeuchi S., Takahashi S.,
RA   Okuda K.;
RT   "Expression and action of hepatocyte growth factor in bovine
RT   endometrial stromal and epithelial cells in vitro.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB056447; BAB33031.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00130; KR; 2.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
FT   NON_TER       1      1
FT   NON_TER      75     75
SQ   SEQUENCE   75 AA;  8831 MW;  829EEEFCC49701B1 CRC64;

  Query Match             69.3%;  Score 52;  DB 6;  Length 75;
  Best Local Similarity   72.7%;  Pred. No. 0.31;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db       4 RNPRGEEGGPW 14


RESULT 13
Q9N1B8
ID   Q9N1B8      PRELIMINARY;      PRT;   109 AA.
AC   Q9N1B8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ENDOMETRIUM FROM DAY 1 CYCLIC UTERUS;
RX   MEDLINE=20297031; PubMed=10819791;
RA   Chen C., Spencer T.E., Bazer F.W.;
RT   "Expression of hepatocyte growth factor and its receptor c-met in the
RT   ovine uterus.";
RL   Biol. Reprod. 62:1844-1850(2000).
DR   EMBL; AF213397; AAF25945.1; -.
DR   HSSP; P14210; 2HGF.
DR   InterPro; IPR000001; Kringle.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER       1      1
FT   NON_TER     109    109
SQ   SEQUENCE   109 AA;  12501 MW;  1F88FE5DBC0D4A5D CRC64;

  Query Match             69.3%;  Score 52;  DB 6;  Length 109;
  Best Local Similarity   72.7%;  Pred. No. 0.46;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db      67 RNPRGEEGGPW 77


RESULT 14
Q9BYM0
ID   Q9BYM0      PRELIMINARY;      PRT;   208 AA.
AC   Q9BYM0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91369928; PubMed=1832556;
RA   Miyazawa K., Kitamura A., Kitamura N.;
RT   "Structural organization and the transcription initiation site of the
RT   human hepatocyte growth factor gene.";
RL   Biochemistry 30:9170-9176(1991).
DR   EMBL; M75971; AAG53459.1; -.
DR   EMBL; M75967; AAG53459.1; JOINED.
DR   EMBL; M75966; AAG53459.1; JOINED.
DR   EMBL; M75968; AAG53459.1; JOINED.
DR   EMBL; M75969; AAG53459.1; JOINED.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR003966; Prothrombin.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23931 MW;  AE9C50DE5A86B37B CRC64;

  Query Match             69.3%;  Score 52;  DB 4;  Length 208;
  Best Local Similarity   72.7%;  Pred. No. 0.92;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     178 RNPRGEEGGPW 188


RESULT 15
Q13494
ID   Q13494      PRELIMINARY;      PRT;   210 AA.
AC   Q13494;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HGF AGONIST/ANTAGONIST.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96278713; PubMed=8662798;
RA   Cioce V., Csaky K.G., Chan A.M.L., Bottaro D.P., Taylor W.G.,
RA   Jensen R., Aaronson S.A., Rubin J.S.;
RT   "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
RT   HGF/scatter factor variant with partial agonist/antagonist activity.";
RL   J. Biol. Chem. 271:13110-13115(1996).
DR   EMBL; U46010; AAC50539.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
SQ   SEQUENCE   210 AA;  24116 MW;  94A6EE9C50DE5A86 CRC64;

  Query Match             69.3%;  Score 52;  DB 4;  Length 210;
  Best Local Similarity   72.7%;  Pred. No. 0.93;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     178 RNPRGEEGGPW 188


Search completed: November  8, 2002, 09:36:10
Job time : 29.6667 secs
```

5

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:36 ; Search time 18 Seconds
(without alignments)
64.060 Million cell updates/sec

Title: US-09-657-431-5
Perfect score: 75
Sequence: 1 RNPDGDVGGPWK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	70	93.3	810	1	PLHU	plasmin (EC 3.4.21
2	70	93.3	810	2	B30848	plasmin (EC 3.4.21
3	64	85.3	455	2	A61545	plasmin (EC 3.4.21
4	64	85.3	460	2	B61545	plasmin (EC 3.4.21
5	64	85.3	711	1	A47136	macrophage-stimula
6	64	85.3	812	1	PLMS	plasmin (EC 3.4.21
7	64	85.3	812	1	PLBO	plasmin (EC 3.4.21
8	63	84.0	4548	1	S00657	apoprotein(a) (EC
9	60	80.0	810	2	I46260	plasmin (EC 3.4.21
10	57	76.0	790	1	PLPG	plasmin (EC 3.4.21
11	56	74.7	716	1	A40332	macrophage-stimula
12	56	74.7	716	1	JC5061	macrophage-stimula
13	55	73.3	625	1	TBBO	thrombin (EC 3.4.2
14	53	70.7	710	1	I51283	hepatocyte growth
15	52	69.3	411	2	I51285	hepatocyte growth
16	52	69.3	728	1	JH0579	hepatocyte growth
17	52	69.3	728	1	A35644	hepatocyte growth
18	52	69.3	728	1	A60185	hepatocyte growth
19	50	66.7	169	2	A40522	plasmin (EC 3.4.21
20	48	64.0	336	2	S33879	plasmin precursor
21	48	64.0	508	2	F70662	probable plcC prot
22	48	64.0	617	2	S10511	thrombin (EC 3.4.2
23	47.5	63.3	961	2	T49228	plasma membrane H+
24	47	62.7	622	1	TBHU	thrombin (EC 3.4.2
25	46	61.3	260	2	H70134	flagellar motor ro
26	46	61.3	312	2	T46255	hypothetical prote
27	46	61.3	618	2	A35827	thrombin (EC 3.4.2
28	46	61.3	943	2	B45082	neurotrophic recep
29	44	58.7	194	2	T10851	y4hQ protein - Rhi
30	43.5	58.0	1420	2	A32869	apolipoprotein(a)
31	43.5	58.0	2869	2	T18518	apolipoprotein(a)
32	43	57.3	183	2	S05569	signaling protein
33	43	57.3	183	2	B90643	AmpD protein [impo
34	43	57.3	183	2	B85494	regulates ampC [im
35	43	57.3	187	2	S26139	signaling protein
36	43	57.3	187	2	A48901	signaling protein
37	43	57.3	187	2	AI0520	AmpD protein (anhy
38	43	57.3	310	2	I50696	collagen alpha 1(I
39	43	57.3	377	2	H70604	probable acyl-coA
40	43	57.3	534	2	F82615	methyltransferase
41	42	56.0	152	2	E87272	hypothetical prote
42	42	56.0	206	2	JC7320	K562 cell-derived
43	41.5	55.3	559	1	A35029	t-plasminogen acti
44	41.5	55.3	559	1	A29941	t-plasminogen acti
45	41.5	55.3	562	1	UKHUT	t-plasminogen acti

ALIGNMENTS

RESULT 1
PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N;Alternate names: plasminogen precursor [misnomer]
N;Contains: angiostatin; microplasmin; plasminogen
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C;Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the f
A;Reference number: A35229; MUID:90202879
A;Accession: A35229
A;Molecule type: DNA
A;Residues: 1-810 <PET>
A;Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A;Experimental source: leukocyte; lung fibroblast
R;Malgaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A;Title: Definition of the transcription initiation site of human plasminogen gene in
A;Reference number: I52242; MUID:91097523
A;Accession: I52242
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <MAL1>
A;Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613
R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human
A;Reference number: A26646; MUID:87162490
A;Accession: A26646
A;Molecule type: mRNA
A;Residues: 1-471,'D',473-810 <FOR>
A;Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
A;Experimental source: liver
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human a
A;Reference number: I45961; MUID:85023311
A;Accession: I62738
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 292-471,'D',473-810 <MAL2>
A;Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A;Accession: I84609
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 367-419 <MAL3>
A;Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg,
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of huma

A;Reference number: S03735; MUID:81212097
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71,'E',73-76 <BRU>
R;Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: protein
A;Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R;Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A;Title: Primary structure of the B-chain of human plasmin.
A;Reference number: A04627; MUID:77225245
A;Accession: A04627
A;Molecule type: protein
A;Residues: 581-810 <WI1>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla

A;Reference number: A04625; MUID:75093329
A;Accession: A04625
A;Molecule type: protein
A;Residues: 20-50,'Q',51-71,'E',73-85,87-100 <WI2>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A;Reference number: A04626; MUID:76043692
A;Accession: A04626
A;Molecule type: protein
A;Residues: 483-507,'E',509-604 <WI3>
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A;Title: The primary structure of human plasminogen. II. The histidine loop of human pla
A;Reference number: A92125; MUID:73149248
A;Contents: annotation; active site
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A;Reference number: A92048; MUID:69234739
A;Contents: annotation; active site
R;Trexler, M.; Vali, Z.; Patthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A;Reference number: A92382; MUID:82213905
A;Contents: annotation; omega-aminocarboxylic acid binding sites
R;Vali, Z.; Patthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A;Reference number: A92458; MUID:85054794
A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A;Title: Kringle domains of human angiostatin. Characterization of the anti-proliferativ
A;Reference number: A58811; MUID:97067211
A;Contents: annotation
R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A;Title: Generation of an angiostatin-like fragment from plasminogen by stromelysin-1 (M
A;Reference number: A58812; MUID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR
A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PMK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.
A;Reference number: A58819; MUID:92031502
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plas
A;Reference number: A58818; MUID:92031503
A;Contents: annotation
R;de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook,
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at
A;Reference number: A39483; MUID:92118803
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmi
A;Reference number: A58817; MUID:94237158
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH
d PIR:FGHUGB).
C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately
rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibit
C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial c
C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiostatin.
ting solid tumors.
C;Genetics:
A;Gene: GDB:PLG
A;Cross-references: GDB:119498; OMIM:173350
A;Map position: 6q26-6q27
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1;
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;20-96/Domain: activation peptide #status experimental <APT>

```
F;79-466/Product: angiostatin #status experimental <AST>
F;97-580,581-810/Product: plasmin #status experimental <MAT>
F;97-580/Domain: plasmin chain A #status experimental <CHA>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;550-580,581-810/Product: microplasmin #status experimental <MMT>

  Query Match             93.3%;  Score 70;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 0.0018;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| |
Db    532 RNPDGDVGGPW 542


RESULT 2
B30848
plasmin (EC 3.4.21.7) precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C;Accession: B32869; B30848
R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A;Reference number: A32869; MUID:89174660
A;Accession: B32869
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-810 <TOM>
A;Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-9/Domain: signal sequence #status predicted <SIG>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;581-803/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
 bonds: #status predicted
F;622,665,760/Active site: His, Asp, Ser #status predicted

  Query Match             93.3%;  Score 70;  DB 2;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 0.0018;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| |
Db    532 RNPDGDVGGPW 542


RESULT 3
A61545
plasmin (EC 3.4.21.7) precursor - horse (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Equus caballus (domestic horse)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: A61545; S17527
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: A61545
A;Molecule type: protein
A;Residues: 1-33;34-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A;Title: Complete amino acid sequence of equine miniplasminogen.
A;Reference number: S17527; MUID:92052077
A;Accession: S17527
A;Molecule type: protein
A;Residues: 118-455 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F;1-33/Domain: activation peptide (fragment) #status experimental <APT>
F;34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F;37-114/Domain: kringle homology <KR4>
F;118-455/Product: miniplasminogen #status experimental <MIN>
F;126-205/Domain: kringle homology <KR5>
F;226-455/Domain: plasmin chain B #status experimental <BCH>
F;226-448/Domain: trypsin homology <TRY>
F;267,310,405/Active site: His, Asp, Ser #status predicted

  Query Match             85.3%;  Score 64;  DB 2;  Length 455;
  Best Local Similarity   90.9%;  Pred. No. 0.0095;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db    177 RNPDGDVNGPW 187


RESULT 4
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37;38-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; MUID:93149995
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
F;1-37/Domain: activation peptide (fragment) #status experimental <APT>
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F;41-118/Domain: kringle homology <KR4>
F;118-460/Product: miniplasminogen #status experimental <MIN>
F;132-211/Domain: kringle homology <KR5>
F;226-460/Domain: plasmin chain B #status experimental <BCH>
F;231-453/Domain: trypsin homology <TRY>
F;272,315,410/Active site: His, Asp, Ser #status predicted

  Query Match             85.3%;  Score 64;  DB 2;  Length 460;
  Best Local Similarity   90.9%;  Pred. No. 0.0096;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db    183 RNPDGDVNGPW 193


RESULT 5
A47136
macrophage-stimulating protein 1 precursor - human
```

```
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C;Accession: A40331; B40331; A47136; A61395
R;Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A;Title: Characterization of the DNF15S2 locus on human chromosome 3: identification of
A;Reference number: A40331; MUID:92002016
A;Accession: A40331
A;Molecule type: DNA
A;Residues: 1-711 <HA1>
A;Cross-references: GB:M74179
A;Accession: B40331
A;Molecule type: mRNA
A;Residues: 1-711 <HA2>
A;Cross-references: GB:M74178; NID:g183976; PIDN:AAA50165.1; PID:g183977
R;Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A;Title: Cloning, sequencing, and expression of human macrophage stimulating protein (MS
A;Reference number: A47136; MUID:93340141
A;Accession: A47136
A;Molecule type: mRNA
A;Residues: 1-12,'C',14-622,'F',624-711 <YOS>
A;Cross-references: GB:L11924; NID:g398037; PIDN:AAA59872.1; PID:g398038
A;Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequence
R;Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A;Title: Macrophage stimulating protein: purification, partial amino acid sequence, and
A;Reference number: A61395; MUID:91217635
A;Accession: A61395
A;Molecule type: protein
A;Residues: 230-247;288-291,'E',293-295,'X',297-301,'X',303,'E',305,'EX',308-310;326-331
A;Experimental source: plasma
C;Genetics:
A;Gene: GDB:MST1; D3F15S2; DNF15S2; HGFL
A;Cross-references: GDB:128833; OMIM:142408
A;Map position: 3p21-3p21.3
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F;19-483/Domain: alpha chain #status predicted <ACH>
F;110-186/Domain: kringle homology <KR1>
F;191-268/Domain: kringle homology <KR2>
F;283-361/Domain: kringle homology <KR3>
F;370-448/Domain: kringle homology <KR4>
F;484-711/Domain: beta chain #status predicted <BCH>
F;484-704/Domain: trypsin homology <TRY>
F;56-78,60-66,110-186,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355,37
F;72,296,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             85.3%;  Score 64;  DB 1;  Length 711;
  Best Local Similarity   90.9%;  Pred. No. 0.015;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| ||||
Db    158 RNPDGDPGGPW 168


RESULT 6
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiostatin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of t
A;Reference number: A38514; MUID:91184812
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DEG>
A;Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LI2>
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, both plas
mediately after dissociation from the clot. In the presence of the inhibitor, the act
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiostatin.
eful in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiostatin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
 bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

  Query Match             85.3%;  Score 64;  DB 1;  Length 812;
  Best Local Similarity   90.9%;  Pred. No. 0.018;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| |||
Db    532 RNPDGDVNGPW 542


RESULT 7
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N;Alternate names: plasminogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C;Accession: S45046; A25835; I45961; S03736
R;Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A;Description: Cloning and characterizatin of the bovine plasminogen cDNA.
A;Reference number: S45046
A;Accession: S45046
A;Molecule type: mRNA
A;Residues: 1-812 <BER>
A;Cross-references: EMBL:X79402; NID:g494962; PIDN:CAA55939.1; PID:g494963
A;Experimental source: liver
A;Note: it is uncertain whether Met-1 or Met-8 is the initiator
R;Schaller, J.; Moser, P.W.; Dannegger-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; R
```

Eur. J. Biochem. 149, 267-278, 1985
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasm
A;Reference number: A25835; MUID:85203906
A;Accession: A25835
A;Molecule type: protein
A;Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A;Reference number: I45961; MUID:85023311
A;Accession: I45961
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 706-743,'R',745-812 <MAL>
A;Cross-references: GB:K02935; NID:g163551; PIDN:AAA30714.1; PID:g163552
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A;Reference number: S03735; MUID:81212097
A;Accession: S03736
A;Molecule type: protein
A;Residues: 27-83 <BRU>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma;
F;1-26/Domain: signal sequence #status predicted <SIG>
F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F;27-812/Product: plasminogen #status experimental <PRO>
F;27-103/Domain: activation peptide #status experimental <APT>
F;104-583,584-812/Product: plasmin #status experimental <MAT>
F;104-583/Domain: plasmin chain A #status experimental <ACH>
F;110-188/Domain: kringle homology <KR1>
F;192-269/Domain: kringle homology <KR2>
F;282-359/Domain: kringle homology <KR3>
F;384-461/Domain: kringle homology <KR4>
F;485-564/Domain: kringle homology <KR5>
F;584-812/Domain: plasmin chain B #status experimental <BCH>
F;584-805/Domain: trypsin homology <TRY>
F;56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,33
bonds: #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 85.3%; Score 64; DB 1; Length 812;
Best Local Similarity 90.9%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db    536 RNPDGDVNGPW 546
```

RESULT 8
S00657
apoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N;Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-2000
C;Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R;McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scan
Nature 330, 132-137, 1987
A;Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A;Reference number: S00657; MUID:88039109
A;Accession: S00657
A;Molecule type: mRNA
A;Residues: 1-4548 <MCL>
A;Cross-references: GB:X06290; EMBL:X06696; NID:g28619; PIDN:CAA29618.1; PID:g28620
R;Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A;Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous
A;Reference number: A28017; MUID:87204109
A;Accession: A28017
A;Molecule type: protein
A;Residues: 20-21,'P',23-34;177-179,'N',181-186,'T',188-196,'DKG',200;292-314,'W',316
X',4396-4401 <EAT>
R;Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A;Title: 5' control regions of the apolipoprotein(a) gene and members of the related
A;Reference number: A47277; MUID:93165698
A;Accession: A47277
A;Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RES>
A;Cross-references: GB:L07899; NID:g967973; PID:g967974
R;Malgaretti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sa
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A;Title: Characterization by yeast artificial chromosome cloning of the linked apolip
A;Reference number: A47233; MUID:93087573
A;Accession: I60906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE2>
A;Cross-references: GB:M90078; NID:g178786; PIDN:AAA35547.1; PID:g553188
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: A47233
A;Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE5>
A;Cross-references: GB:M90079; NID:g178784; PIDN:AAA35546.1; PID:g553187
R;Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A;Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated
A;Reference number: I52415; MUID:92207924
A;Accession: I52415
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE3>
A;Cross-references: GB:M86877; NID:g178780; PIDN:AAB49909.1; PID:g553185
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: I65286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE4>
A;Cross-references: GB:M86878; NID:g178782; PIDN:AAA51749.1; PID:g553186
C;Genetics:
A;Gene: GDB:LPA
A;Cross-references: GDB:120699; OMIM:152200
A;Map position: 6q26-6q27
A;Note: several genes closely linked on chromosome 6 are identical in the first codin
rs of kringle repeats
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F;28-105/Domain: kringle homology <KR1>
F;142-219/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;370-447/Domain: kringle homology <KR4>
F;484-561/Domain: kringle homology <KR5>
F;598-675/Domain: kringle homology <KR6>
F;712-789/Domain: kringle homology <KR7>
F;826-903/Domain: kringle homology <KR8>
F;940-1017/Domain: kringle homology <KR9>
F;1054-1131/Domain: kringle homology <KR10>
F;1168-1245/Domain: kringle homology <KR11>
F;1282-1359/Domain: kringle homology <KR12>
F;1396-1473/Domain: kringle homology <KR13>
F;1510-1587/Domain: kringle homology <KR14>
F;1624-1701/Domain: kringle homology <KR15>
F;1738-1815/Domain: kringle homology <KR16>
F;1852-1929/Domain: kringle homology <KR17>
F;1966-2043/Domain: kringle homology <KR18>

```
F;2080-2157/Domain: kringle homology <KR19>
F;2194-2271/Domain: kringle homology <KR20>
F;2308-2385/Domain: kringle homology <KR21>
F;2422-2499/Domain: kringle homology <KR22>
F;2536-2613/Domain: kringle homology <KR23>
F;2650-2727/Domain: kringle homology <KR24>
F;2764-2841/Domain: kringle homology <KR25>
F;2878-2955/Domain: kringle homology <KR26>
F;2992-3069/Domain: kringle homology <KR27>
F;3106-3183/Domain: kringle homology <KR28>
F;3220-3297/Domain: kringle homology <KR29>
F;3334-3411/Domain: kringle homology <KR30>
F;3448-3525/Domain: kringle homology <KR31>
F;3562-3639/Domain: kringle homology <KR32>
F;3676-3753/Domain: kringle homology <KR33>
F;3782-3859/Domain: kringle homology <KR34>
F;3896-3973/Domain: kringle homology <KR35>
F;4010-4087/Domain: kringle homology <KR36>
F;4124-4201/Domain: kringle homology <KR37>
F;4228-4307/Domain: kringle homology <KR38>
F;4328-4541/Domain: trypsin homology <TRY>

  Query Match             84.0%;  Score 63;  DB 1;  Length 4548;
  Best Local Similarity   81.8%;  Pred. No. 0.16;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| |||
Db   4279 RNPDGDINGPW 4289


RESULT 9
I46260
plasmin (EC 3.4.21.7) precursor - western European hedgehog
C;Species: Erinaceus europaeus (western European hedgehog)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46260
R;Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A;Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprot
A;Reference number: I46259; MUID:96025778
A;Accession: I46260
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-810 <LAW>
A;Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: hydrolase; serine proteinase
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;379-456/Domain: kringle homology <KR4>
F;482-561/Domain: kringle homology <KR5>
F;582-803/Domain: trypsin homology <TRY>

  Query Match             80.0%;  Score 60;  DB 2;  Length 810;
  Best Local Similarity   81.8%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| |||
Db    533 RNPDGDANGPW 543


RESULT 10
PLPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: S03733; S03737; A25834
R;Schaller, J.; Marti, T.; Roesselet, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A;Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the
A;Reference number: S03733
A;Accession: S03733
A;Molecule type: protein
A;Residues: 1-560 <SCH>
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg,
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of huma
A;Reference number: S03735; MUID:81212097
A;Accession: S03737
A;Molecule type: protein
A;Residues: 1-57 <BRU>
R;Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A;Title: Determination of the complete amino-acid sequence of porcine miniplasminogen
A;Reference number: A25834; MUID:85203907
A;Accession: A25834
A;Molecule type: protein
A;Residues: 450-790 <MAR>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine pr
F;1-790/Product: plasminogen #status predicted <PRO>
F;1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F;1-77/Domain: activation peptide #status predicted <APT>
F;78-560/Product: plasmin chain A #status predicted <ACH>
F;84-162/Domain: kringle homology <KR1>
F;166-243/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;358-435/Domain: kringle homology <KR4>
F;450-790/Product: miniplasminogen #status experimental <MIN>
F;461-540/Domain: kringle homology <KR5>
F;561-790/Product: plasmin chain B #status experimental <BCH>
F;561-783/Domain: trypsin homology <TRY>
F;30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,
bonds: #status predicted
F;602,645,740/Active site: His, Asp, Ser #status predicted

  Query Match             76.0%;  Score 57;  DB 1;  Length 790;
  Best Local Similarity   81.8%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          |||||| |||
Db    512 RNPDGDDNGPW 522


RESULT 11
A40332
macrophage-stimulating protein 1 precursor - mouse
N;Alternate names: hepatocyte growth factor-like protein
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C;Accession: A40332; B40332
R;Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
Biochemistry 30, 9781-9791, 1991
A;Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth f
A;Reference number: A40332; MUID:92002017
A;Accession: A40332
A;Molecule type: DNA
A;Residues: 1-716 <DEG>
A;Cross-references: GB:M74180; NID:g193831; PIDN:AAA50166.1; PID:g193832
A;Accession: B40332
A;Molecule type: mRNA
A;Residues: 1-18,'P',20-716 <DEG2>
A;Cross-references: GB:M74181; NID:g193833; PIDN:AAA50167.1; PID:g193834
C;Genetics:
A;Introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1;
```

C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle
F;1-31/Domain: signal sequence #status predicted <SIG>
F;19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
F;19-483/Domain: alpha chain #status experimental <ACH>
F;110-186/Domain: kringle homology <KR1>
F;191-268/Domain: kringle homology <KR2>
F;292-370/Domain: kringle homology <KR3>
F;379-457/Domain: kringle homology <KR4>
F;484-711/Domain: beta chain #status experimental <BCH>
F;489-709/Domain: trypsin homology <TRY>
F;72,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.7%; Score 56; DB 1; Length 716;
Best Local Similarity 81.8%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
|||||| |||
Db 429 RNPDGDSHGPW 439

RESULT 12
JC5061
macrophage-stimulating protein 1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C;Accession: JC5061
R;Ohshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A;Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A;Reference number: JC5061; MUID:97011126
A;Accession: JC5061
A;Molecule type: mRNA
A;Residues: 1-716 <OHS>
A;Cross-references: EMBL:X95096; NID:g1669718; PIDN:CAA64473.1; PID:g1669719
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F;32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F;110-186/Domain: kringle homology <KRI1>
F;191-268/Domain: kringle homology <KRI2>
F;292-370/Domain: kringle homology <KRI3>
F;379-457/Domain: kringle homology <KRI4>
F;489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F;489-709/Domain: trypsin homology <TRY>
F;72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.7%; Score 56; DB 1; Length 716;
Best Local Similarity 81.8%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPW 11
|||||| |||
Db 429 RNPDGDSHGPW 439

RESULT 13
TBBO
thrombin (EC 3.4.21.5) precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
C;Accession: S02537; A00915; A37552; I46045; S67518
R;Irwin, D.M.; Robertson, K.A.; MacGillivray, R.T.A.
J. Mol. Biol. 200, 31-45, 1988
A;Title: Structure and evolution of the bovine prothrombin gene.
A;Reference number: S02537; MUID:88245190
A;Accession: S02537
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-625 <IRW>
R;MacGillivray, R.T.A.; Davie, E.W.
Biochemistry 23, 1626-1634, 1984
A;Title: Characterization of bovine prothrombin mRNA and its translation product.
A;Reference number: A00915; MUID:84203525
A;Accession: A00915
A;Molecule type: mRNA
A;Residues: 1-230,'H',232-625 <MAC>
A;Note: 600-Asn was also found
R;Magnusson, S.; Sottrup-Jensen, L.; Petersen, T.E.; Claeys, H.
in Boerhaave Symposium on Prothrombin and Related Coagulation Factors, Hemker, H.C.,
A;Reference number: A37552
A;Accession: A37552
A;Molecule type: protein
A;Residues: 44-287,'N',289-352,'E',354,'Q',356-548,'ND',551-599,'N',601-625 <MAG>
A;Note: the evidence for 231-Ser is strong
A;Note: disulfide bonds and carbohydrate binding sites were determined
R;Park, C.H.; Tulinsky, A.
Biochemistry 25, 3977-3982, 1986
A;Title: Three-dimensional structure of the kringle sequence: structure of prothrombi
A;Reference number: A37553; MUID:86296631
A;Contents: annotation; residues 44-317, X-ray crystallography, 2.8 angstroms
R;Irwin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.
Biochemistry 24, 6854-6861, 1985
A;Title: Characterization of the bovine prothrombin gene.
A;Reference number: A37554; MUID:86077733
A;Contents: annotation; gene structure
R;MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980
A;Title: Cloning and analysis of a cDNA coding for bovine prothrombin.
A;Reference number: I46045; MUID:81054926
A;Accession: I46045
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 466-599,'N',601-625 <MA2>
A;Cross-references: EMBL:V00135; NID:g772; PIDN:CAA23451.1; PID:g808945
R;Pejler, G.; Karlstroem, A.R.; Berg, L.
Eur. J. Biochem. 227, 102-107, 1995
A;Title: Identification of the proteolytic thrombin fragments formed after cleavage w
A;Reference number: S67518; MUID:95154277
A;Accession: S67518
A;Status: preliminary
A;Molecule type: protein
A;Residues: 318-325;333-338,'X',340;367-374;481-484,'X',486-488;515-522 <PEJ>
C;Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fi
C;Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi
tivation peptide and cleaves the remaining part into light and heavy chains. The acti
.
C;Comment: Thrombin can cleave the amino-terminal activation peptide 1 from prothromb
C;Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carb
ent interaction with the negatively charged phospholipid membrane surface.
C;Comment: The prothrombin precursor is synthesized in the liver.
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; duplication; gl
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-625/Product: prothrombin #status experimental <MPT>
F;44-199/Domain: activation peptide 1 #status experimental <FR1>
F;109-187/Domain: kringle homology <KR1>
F;200-317/Domain: activation peptide 2 #status experimental <FR2>
F;214-292/Domain: kringle homology <KR2>
F;318-366/Product: thrombin light chain #status experimental <LCH>
F;367-625/Product: thrombin heavy chain #status experimental <HCH>
F;367-616/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;61-66,91-104,109-187,130-170,158-182,214-292,235-275,263-287,339-485,394-410,539-55
F;120,144,419/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;409,465,571/Active site: His, Asp, Ser #status experimental

Query Match 73.3%; Score 55; DB 1; Length 625;
Best Local Similarity 72.7%; Pred. No. 0.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPW 11
           ||||| : |||
Db     159 RNPDGSITGPW 169
```

RESULT 14
I51283
hepatocyte growth factor precursor - clawed frog
N;Alternate names: hepapoietin A; scatter factor
C;Species: Xenopus sp. (clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I51283
R;Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A;Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A;Reference number: I51283; MUID:95267690
A;Accession: I51283
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-710 <NAK>
A;Cross-references: GB:S77422; NID:g998932; PIDN:AAB34354.1; PID:g998933
A;Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleoti
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F;42-477,478-709/Product: hepatocyte growth factor #status predicted <MAT>
F;42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;115-193/Domain: kringle homology <KR1>
F;198-275/Domain: kringle homology <KR2>
F;289-367/Domain: kringle homology <KR3>
F;375-453/Domain: kringle homology <KR4>
F;478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;478-700/Domain: trypsin homology <TRY>
F;52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre
F;470-588/Disulfide bonds: #status predicted

  Query Match 70.7%; Score 53; DB 1; Length 710;
  Best Local Similarity 72.7%; Pred. No. 0.97;
  Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     165 RNPKGEEGGPW 175
```

RESULT 15
I51285
hepatocyte growth factor/scatter factor - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51285
R;Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gherard
Development 121, 813-824, 1995
A;Title: A role for HGF/SF in neural induction and its expression in Hensen's node durin
A;Reference number: I51285; MUID:95237013
A;Accession: I51285
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-411 <STR>
A;Cross-references: GB:S77480; NID:g998675; PID:g998676
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F;124-197/Domain: kringle homology <KRG>
F;202-279/Domain: kringle homology <KR2>
F;296-374/Domain: kringle homology <KR3>

  Query Match 69.3%; Score 52; DB 2; Length 411;
  Best Local Similarity 72.7%; Pred. No. 0.79;
  Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     169 RNPRGEEGGPW 179
```

Search completed: November 8, 2002, 09:36:44
Job time : 19 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:33:21 ; Search time 9.33333 Seconds
(without alignments)
49.782 Million cell updates/sec

Title: US-09-657-431-5
Perfect score: 75
Sequence: 1 RNPDGDVGGPWK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	70	93.3	810	1	PLMN_HUMAN	P00747 homo sapien
2	70	93.3	810	1	PLMN_MACMU	P12545 macaca mula
3	64	85.3	333	1	PLMN_CANFA	P80009 canis famil
4	64	85.3	338	1	PLMN_HORSE	P80010 equus cabal
5	64	85.3	343	1	PLMN_SHEEP	P81286 ovis aries
6	64	85.3	711	1	HGFL_HUMAN	P26927 homo sapien
7	64	85.3	812	1	PLMN_BOVIN	P06868 bos taurus
8	64	85.3	812	1	PLMN_MOUSE	P20918 mus musculu
9	63	84.0	4548	1	APOA_HUMAN	P08519 homo sapien
10	60	80.0	810	1	PLMN_ERIEU	Q29485 erinaceus e
11	57	76.0	790	1	PLMN_PIG	P06867 sus scrofa
12	56	74.7	716	1	HGFL_MOUSE	P26928 mus musculu
13	55	73.3	625	1	THRB_BOVIN	P00735 bos taurus
14	52	69.3	728	1	HGF_HUMAN	P14210 homo sapien
15	52	69.3	728	1	HGF_MOUSE	Q08048 mus musculu
16	52	69.3	728	1	HGF_RAT	P17945 rattus norv
17	50	66.7	169	1	PLMN_RAT	Q01177 rattus norv
18	48	64.0	325	1	PLMN_PETMA	P33574 petromyzon
19	48	64.0	517	1	PHLC_MYCTU	P95245 mycobacteri
20	48	64.0	617	1	THRB_RAT	P18292 rattus norv
21	47.5	63.3	961	1	PMA7_ARATH	Q9ly32 arabidopsis
22	47	62.7	622	1	THRB_HUMAN	P00734 homo sapien
23	46	61.3	618	1	THRB_MOUSE	P19221 mus musculu
24	46	61.3	943	1	ROR2_HUMAN	Q01974 homo sapien
25	46	61.3	944	1	ROR2_MOUSE	Q9z138 mus musculu
26	44	58.7	194	1	Y4HQ_RHISN	P50361 rhizobium s
27	43.5	58.0	1420	1	APOA_MACMU	P14417 macaca mula
28	43	57.3	183	1	AMPD_ECOLI	P13016 escherichia
29	43	57.3	187	1	AMPD_CITFR	P82974 citrobacter
30	43	57.3	187	1	AMPD_ENTCL	P82973 enterobacte
31	43	57.3	187	1	AMPD_SALTY	P30013 salmonella
32	41.5	55.3	559	1	TPA_MOUSE	P11214 mus musculu
33	41.5	55.3	559	1	TPA_RAT	P19637 rattus norv
34	41.5	55.3	562	1	TPA_HUMAN	P00750 homo sapien
35	41.5	55.3	566	1	TPA_BOVIN	Q28198 bos taurus
36	41	54.7	1089	1	UBP6_HUMAN	P35125 homo sapien
37	41	54.7	1748	1	POLR_ELV	P35928 erysimum la
38	41	54.7	1822	1	ZAP3_HUMAN	P49750 homo sapien
39	40	53.3	285	1	PNMT_RAT	P10937 rattus norv
40	40	53.3	295	1	PNMT_MOUSE	P40935 mus musculu
41	40	53.3	1262	1	CA13_CHICK	P12105 gallus gall
42	39	52.0	102	1	SGP3_CHRVI	O52055 chromatium
43	39	52.0	209	1	VP32_BPAPS	Q9t1r6 bacteriopha
44	39	52.0	283	1	CC19_CAEEL	P18835 caenorhabdi
45	39	52.0	399	1	IRTF_MOUSE	Q61179 mus musculu

ALIGNMENTS

RESULT 1
PLMN_HUMAN
ID PLMN_HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;

RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vali Z., Patthy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vali Z., Patthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RT are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RT plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE 268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevich R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;
RL Submitted (JUN-1995) to the PDB data bank.
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.
RX MEDLINE=91095410; PubMed=1986355;
RA Ichinose A., Espling E.S., Takamatsu J., Saito H., Shinmyozu K.,

```
  Query Match             93.3%;  Score 70;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db    532 RNPDGDVGGPW 542


RESULT 2
PLMN_MACMU
ID   PLMN_MACMU     STANDARD;      PRT;   810 AA.
AC   P12545;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89174660; PubMed=2925643;
RA   Tomlinson J.E., McLean J.W., Lawn R.M.;
RT   "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT   synthesis.";
RL   J. Biol. Chem. 264:5957-5965(1989).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-580, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04697; AAA36901.1; -.
DR   PIR; B30848; B30848.
DR   PIR; B32869; B32869.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL       1     19
FT   CHAIN       20    810       PLASMINOGEN.
FT   CHAIN       20    580       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE     20     96       ACTIVATION PEPTIDE.
FT   CHAIN       97    580       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN      581    810       PLASMIN LIGHT CHAIN B.
FT   DOMAIN     103    181       KRINGLE 1.
FT   DOMAIN     184    262       KRINGLE 2.
FT   DOMAIN     275    352       KRINGLE 3.
FT   DOMAIN     377    454       KRINGLE 4.
FT   DOMAIN     481    560       KRINGLE 5.
FT   DOMAIN     581    810       SERINE PROTEASE.
FT   ACT_SITE   622    622       CHARGE RELAY SYSTEM.
FT   ACT_SITE   665    665       CHARGE RELAY SYSTEM.
FT   ACT_SITE   760    760       CHARGE RELAY SYSTEM.
FT   BINDING    136    136       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    158    158       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    172    172       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    432    432       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    445    445       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    134    134       FIBRIN.
FT   BINDING    136    136       FIBRIN.
FT   DISULFID    49     73       BY SIMILARITY.
FT   DISULFID    53     61       BY SIMILARITY.
FT   DISULFID   103    181       BY SIMILARITY.
FT   DISULFID   124    164       BY SIMILARITY.
FT   DISULFID   152    176       BY SIMILARITY.
FT   DISULFID   185    262       BY SIMILARITY.
FT   DISULFID   188    316       BY SIMILARITY.
FT   DISULFID   206    245       BY SIMILARITY.
FT   DISULFID   234    257       BY SIMILARITY.
FT   DISULFID   275    352       BY SIMILARITY.
FT   DISULFID   296    335       BY SIMILARITY.
FT   DISULFID   324    347       BY SIMILARITY.
FT   DISULFID   377    454       BY SIMILARITY.
FT   DISULFID   398    437       BY SIMILARITY.
FT   DISULFID   426    449       BY SIMILARITY.
FT   DISULFID   481    560       BY SIMILARITY.
FT   DISULFID   502    543       BY SIMILARITY.
FT   DISULFID   531    555       BY SIMILARITY.
FT   DISULFID   567    685       BY SIMILARITY.
FT   DISULFID   577    585       BY SIMILARITY.
FT   DISULFID   607    623       BY SIMILARITY.
FT   DISULFID   699    766       BY SIMILARITY.
FT   DISULFID   729    745       BY SIMILARITY.
FT   DISULFID   756    784       BY SIMILARITY.
FT   CARBOHYD   365    365       O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   810 AA;  90255 MW;  A75E1C51A1A0F24A CRC64;

  Query Match             93.3%;  Score 70;  DB 1;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||||||| 
Db    532 RNPDGDVGGPW 542


RESULT 3
PLMN_CANFA
ID   PLMN_CANFA     STANDARD;      PRT;   333 AA.
```

```
AC   P80009;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=90175323; PubMed=2626424;
RA   Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT   "Complete amino acid sequence of canine miniplasminogen.";
RL   Protein Seq. Data Anal. 2:445-450(1989).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF
CC       STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   CHAIN        <1    103       PLASMIN HEAVY CHAIN A.
FT   CHAIN       104    333       PLASMIN LIGHT CHAIN B.
FT   DOMAIN        4     83       KRINGLE 5.
FT   DOMAIN      104    333       SERINE PROTEASE.
FT   DISULFID      4     83       BY SIMILARITY.
FT   DISULFID     25     66       BY SIMILARITY.
FT   DISULFID     54     78       BY SIMILARITY.
FT   DISULFID     90    208       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    100    108       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    130    146       BY SIMILARITY.
FT   DISULFID    222    289       BY SIMILARITY.
FT   DISULFID    252    268       BY SIMILARITY.
FT   DISULFID    279    307       BY SIMILARITY.
FT   ACT_SITE    145    145       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    188    188       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    283    283       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE        152    152       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        186    186       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        264    264       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        277    277       SITE OF SUBSTRATE SPECIFICITY
FT                                (BY SIMILARITY).
SQ   SEQUENCE   333 AA;  36678 MW;  C8C0271B6C6AC8D4 CRC64;
```

```
  Query Match           85.3%;  Score 64;  DB 1;  Length 333;
  Best Local Similarity   90.9%;  Pred. No. 0.003;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||| |||
Db      55 RNPDGDVNGPW 65


RESULT 4
PLMN_HORSE
ID   PLMN_HORSE     STANDARD;      PRT;   338 AA.
AC   P80010;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=92052077; PubMed=1946332;
RA   Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT   "Complete amino acid sequence of equine miniplasminogen.";
RL   Protein Seq. Data Anal. 4:69-74(1991).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
DR   PIR; S17527; S17527.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   CHAIN        <1    108       PLASMIN HEAVY CHAIN A.
FT   CHAIN       109    338       PLASMIN LIGHT CHAIN B.
FT   DOMAIN        9     88       KRINGLE 5.
FT   DOMAIN      109    338       SERINE PROTEASE.
FT   DISULFID      9     88       BY SIMILARITY.
FT   DISULFID     30     71       BY SIMILARITY.
FT   DISULFID     59     83       BY SIMILARITY.
FT   DISULFID     95    213       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    105    113       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    135    151       BY SIMILARITY.
FT   DISULFID    227    294       BY SIMILARITY.
```

```
FT   DISULFID    257    273       BY SIMILARITY.
FT   DISULFID    284    312       BY SIMILARITY.
FT   ACT_SITE    150    150       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    193    193       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    288    288       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE        157    157       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        191    191       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        269    269       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        282    282       SITE OF SUBSTRATE SPECIFICITY
FT                                 (BY SIMILARITY).
SQ   SEQUENCE   338 AA;  37132 MW;  8E9E56B5C5CDBE01 CRC64;

  Query Match             85.3%;  Score 64;  DB 1;  Length 338;
  Best Local Similarity   90.9%;  Pred. No. 0.0031;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db     60 RNPDGDVNGPW 70


RESULT 5
PLMN_SHEEP
ID   PLMN_SHEEP     STANDARD;      PRT;   343 AA.
AC   P81286;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93149995; PubMed=1492092;
RA   Schaller J., Straub C., Kampfer U., Rickli E.E.;
RT   "Complete amino acid sequence of ovine miniplasminogen.";
RL   Protein Seq. Data Anal. 5:21-25(1992).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 2 KRINGLE DOMAINS.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   DOMAIN       <1    140       HEAVY CHAIN A.
FT   DOMAIN      141   >343       LIGHT CHAIN A.
FT   DOMAIN       <1     17       KRINGLE 4.
FT   DOMAIN       41    120       KRINGLE 5.
FT   DOMAIN      114    341       SERINE PROTEASE.
FT   ACT_SITE    181    181       CHARGE RELAY SYSTEM.
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM.
FT   ACT_SITE    319    319       CHARGE RELAY SYSTEM.
FT   NON_TER     343    343
SQ   SEQUENCE   343 AA;  37662 MW;  8DF6EBA92D596EE0 CRC64;

  Query Match             85.3%;  Score 64;  DB 1;  Length 343;
  Best Local Similarity   90.9%;  Pred. No. 0.0031;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db     66 RNPDGDVNGPW 76


RESULT 6
HGFL_HUMAN
ID   HGFL_HUMAN     STANDARD;      PRT;   711 AA.
AC   P26927; Q13350; Q14870;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor-like protein precursor (Macrophage
DE   stimulatory protein) (MSP) (Macrophage stimulating protein).
GN   MST1 OR HGFL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92002016; PubMed=1655021;
RA   Han S., Stuart L.A., Friezner Degen S.J.;
RT   "Characterization of the DNF15S2 locus on human chromosome 3:
RT   identification of a gene coding for four kringle domains with
RT   homology to hepatocyte growth factor.";
RL   Biochemistry 30:9768-9780(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93340141; PubMed=8393443;
RA   Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT   "Cloning, sequencing, and expression of human macrophage stimulating
RT   protein (MSP, MST1) confirms MSP as a member of the family of kringle
RT   proteins and locates the MSP gene on chromosome 3.";
RL   J. Biol. Chem. 268:15461-15468(1993).
CC   -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC       CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC       CONSERVED.
CC   -!- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE
CC       HELD TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M74178; AAA50165.1; -.
DR   EMBL; U37055; AAC50471.1; -.
DR   EMBL; L11924; AAA59872.1; -.
DR   PIR; A40331; A40331.
```

```
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.975; -.
DR   MIM; 142408; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Kringle; Glycoprotein; Serine protease homolog; Signal;
KW   Polymorphism.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    711       HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
FT   DOMAIN       32    109       PAP.
FT   DOMAIN      110    186       KRINGLE 1.
FT   DOMAIN      191    268       KRINGLE 2.
FT   DOMAIN      283    361       KRINGLE 3.
FT   DOMAIN      370    448       KRINGLE 4.
FT   DOMAIN      484    711       SERINE PROTEASE-LIKE.
FT   DISULFID     56     78       BY SIMILARITY.
FT   DISULFID     60     66       BY SIMILARITY.
FT   DISULFID    110    186       BY SIMILARITY.
FT   DISULFID    131    169       BY SIMILARITY.
FT   DISULFID    157    181       BY SIMILARITY.
FT   DISULFID    191    268       BY SIMILARITY.
FT   DISULFID    194    324       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    212    251       BY SIMILARITY.
FT   DISULFID    240    263       BY SIMILARITY.
FT   DISULFID    283    361       BY SIMILARITY.
FT   DISULFID    304    343       BY SIMILARITY.
FT   DISULFID    332    355       BY SIMILARITY.
FT   DISULFID    370    448       BY SIMILARITY.
FT   DISULFID    391    431       BY SIMILARITY.
FT   DISULFID    419    443       BY SIMILARITY.
FT   DISULFID    468    588       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    507    523       BY SIMILARITY.
FT   DISULFID    602    667       BY SIMILARITY.
FT   DISULFID    632    646       BY SIMILARITY.
FT   DISULFID    657    685       BY SIMILARITY.
FT   CARBOHYD     72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      13     13       Y -> C.
FT                                /FTId=VAR_006631.
FT   VARIANT     212    212       C -> F.
FT                                /FTId=VAR_006632.
FT   CONFLICT    623    623       L -> F (IN REF. 2).
SQ   SEQUENCE   711 AA;  80379 MW;  596ED21F180290E4 CRC64;

  Query Match             85.3%;  Score 64;  DB 1;  Length 711;
  Best Local Similarity   90.9%;  Pred. No. 0.0065;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           |||||| ||||
Db     158 RNPDGDPGGPW 168


RESULT 7
PLMN_BOVIN
ID   PLMN_BOVIN     STANDARD;      PRT;   812 AA.
AC   P06868; Q28162;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Berglund L., Andersen M.D., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen cDNA.";
RL   Int. Dairy J. 5:593-603(1995).
RN   [2]
RP   SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=85203906; PubMed=3846532;
RA   Schaller J., Moser P.W., Dannegger-Muller G.A.K., Rosselet S.J.,
RA   Kampfer U., Rickli E.E.;
RT   "Complete amino acid sequence of bovine plasminogen. Comparison with
RT   human plasminogen.";
RL   Eur. J. Biochem. 149:267-278(1985).
RN   [3]
RP   SEQUENCE OF 706-812 FROM N.A.
RX   MEDLINE=85023311; PubMed=6148961;
RA   Malinowski D.P., Sadler J.E., Davie E.W.;
RT   "Characterization of a complementary deoxyribonucleic acid coding for
RT   human and bovine plasminogen.";
RL   Biochemistry 23:4243-4250(1984).
RN   [4]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=88185329; PubMed=3356193;
RA   Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA   Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT   "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT   plasminogen. Species specificity in relation to sialylation and
RT   fucosylation patterns.";
RL   Eur. J. Biochem. 173:57-63(1988).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLLACTOSAMINE AND SIALIC ACID.
CC       O-LINKED GLYCANS CONSIST OF GAL-GALNAC DISACCHARIDE WITH IS
CC       MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X79402; CAA55939.1; -.
DR   EMBL; K02935; AAA30714.1; -.
DR   PIR; A25835; PLBO.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.233; -.
DR   GlycoSuiteDB; P06868; -.
```

```
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL        1     26
FT   CHAIN        27    812       PLASMINOGEN.
FT   CHAIN        27    583       PLASMIN HEAVY CHAIN A.
FT   CHAIN       584    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN      110    188       KRINGLE 1.
FT   DOMAIN      192    269       KRINGLE 2.
FT   DOMAIN      282    359       KRINGLE 3.
FT   DOMAIN      384    461       KRINGLE 4.
FT   DOMAIN      485    564       KRINGLE 5.
FT   DOMAIN      584    812       SERINE PROTEASE.
FT   CARBOHYD    315    315       N-LINKED (GLCNAC. . .).
FT                                /FTId=CAR_000014.
FT   CARBOHYD    365    365       O-LINKED (GALNAC. . .).
FT                                /FTId=CAR_000015.
FT   ACT_SITE    624    624       CHARGE RELAY SYSTEM.
FT   ACT_SITE    667    667       CHARGE RELAY SYSTEM.
FT   ACT_SITE    762    762       CHARGE RELAY SYSTEM.
FT   CONFLICT    335    335       N -> D (IN REF. 2).
FT   CONFLICT    516    516       Q -> H (IN REF. 2).
FT   CONFLICT    555    555       P -> L (IN REF. 2).
FT   CONFLICT    744    744       T -> R (IN REF. 3).
SQ   SEQUENCE   812 AA;  91216 MW;  38A6AA691E220946 CRC64;

  Query Match             85.3%;  Score 64;  DB 1;  Length 812;
  Best Local Similarity   90.9%;  Pred. No. 0.0075;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db    536 RNPDGDVNGPW 546


RESULT 8
PLMN_MOUSE
ID   PLMN_MOUSE     STANDARD;      PRT;   812 AA.
AC   P20918;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN   PLG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91184812; PubMed=2081600;
RA   Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse plasminogen and
RT   localization of the gene to mouse chromosome 17.";
RL   Genomics 8:49-61(1990).
RN   [2]
RP   CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX   MEDLINE=95042728; PubMed=7525077;
RA   O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA   Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT   "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT   suppression of metastases by a Lewis lung carcinoma.";
RL   Cell 79:315-328(1994).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC       NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC       METASTATIC TUMORS IN VIVO.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04766; AAA50168.1; -.
DR   PIR; A38514; A38514.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   MGD; MGI:97620; Plg.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
FT   CHAIN        20    581       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE      20     97       ACTIVATION PEPTIDE.
FT   CHAIN        98    581       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN        98   ?436       ANGIOSTATIN.
FT   CHAIN       582    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN      103    181       KRINGLE 1.
FT   DOMAIN      184    262       KRINGLE 2.
```

```
FT   DOMAIN      275    352       KRINGLE 3.
FT   DOMAIN      377    454       KRINGLE 4.
FT   DOMAIN      481    560       KRINGLE 5.
FT   DOMAIN      582    812       SERINE PROTEASE.
FT   ACT_SITE    624    624       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    667    667       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    762    762       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     49     73       BY SIMILARITY.
FT   DISULFID     53     61       BY SIMILARITY.
FT   DISULFID    103    181       BY SIMILARITY.
FT   DISULFID    124    164       BY SIMILARITY.
FT   DISULFID    152    176       BY SIMILARITY.
FT   DISULFID    185    262       BY SIMILARITY.
FT   DISULFID    188    316       BY SIMILARITY.
FT   DISULFID    206    245       BY SIMILARITY.
FT   DISULFID    234    257       BY SIMILARITY.
FT   DISULFID    275    352       BY SIMILARITY.
FT   DISULFID    296    335       BY SIMILARITY.
FT   DISULFID    324    347       BY SIMILARITY.
FT   DISULFID    377    454       BY SIMILARITY.
FT   DISULFID    398    437       BY SIMILARITY.
FT   DISULFID    426    449       BY SIMILARITY.
FT   DISULFID    481    560       BY SIMILARITY.
FT   DISULFID    502    543       BY SIMILARITY.
FT   DISULFID    531    555       BY SIMILARITY.
FT   DISULFID    568    687       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    578    586       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    609    625       BY SIMILARITY.
FT   DISULFID    701    768       BY SIMILARITY.
FT   DISULFID    731    747       BY SIMILARITY.
FT   DISULFID    758    786       BY SIMILARITY.
SQ   SEQUENCE   812 AA;  90846 MW;  D34A74A4FC2256F8 CRC64;

  Query Match             85.3%;  Score 64;  DB 1;  Length 812;
  Best Local Similarity   90.9%;  Pred. No. 0.0075;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||||| |||
Db    532 RNPDGDVNGPW 542


RESULT 9
APOA_HUMAN
ID   APOA_HUMAN     STANDARD;      PRT;  4548 AA.
AC   P08519;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN   LPA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88039109; PubMed=3670400;
RA   McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA   Fless G.M., Scanu A.M., Lawn R.M.;
RT   "cDNA sequence of human apolipoprotein(a) is homologous to
RT   plasminogen.";
RL   Nature 330:132-137(1987).
RN   [2]
RP   SERINE PROTEASE ACTIVITY.
RX   MEDLINE=90076123; PubMed=2531657;
RA   Salonen E.-M., Jauhiainen M., Zardi L., Vaheri A., Ehnholm C.;
RT   "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT   activity capable of cleaving it.";
RL   EMBO J. 8:4035-4040(1989).
RN   [3]
RP   REVIEW.
RX   MEDLINE=90049223; PubMed=2530631;
RA   Utermann G.;
RT   "The mysteries of lipoprotein(a).";
RL   Science 246:904-910(1989).
RN   [4]
RP   CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX   MEDLINE=21303595; PubMed=11294842;
RA   Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT   "Structural elucidation of the N- and O-glycans of human
RT   apolipoprotein(a): role of o-glycans in conferring protease
RT   resistance.";
RL   J. Biol. Chem. 276:22200-22208(2001).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX   MEDLINE=96217891; PubMed=8642595;
RA   Mikol V., Lograsso P.V., Boettcher B.R.;
RT   "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT   complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RT   acid: existence of novel and expected binding modes.";
RL   J. Mol. Biol. 256:751-761(1996).
RN   [6]
RP   VARIANT ARG-4193.
RX   MEDLINE=95002201; PubMed=7918682;
RA   Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT   "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT   associated with a lysine binding defect in Lp(a).";
RL   Biochim. Biophys. Acta 1227:41-45(1994).
CC   -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC       (Lp(a)). It has serine proteinase activity and is able of
CC       autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC       Lp(a) may be a ligand for megalin/Gp 330.
CC   -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC       decorin.
CC   -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC       structures present in either a mono- or disialylated state. The
CC       O-glycans are mostly (80%) represented by the monosialylated core
CC       type I structure, NeuNAcalpha2-3Galbeta1-3GalNAc, with smaller
CC       amounts of disialylated and non-sialylated O-glycans also
CC       detected.
CC   -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC       naturally occuring proteolytic fragments is correlated with
CC       atherosclerosis. Homology with plasminogen kringles IV and V is
CC       thought to underlie the atherogenicity of the protein, because the
CC       fragments are competing with plasminogen for fibrin(ogen) binding.
CC   -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC       leading to the formation of the so called mini-Lp(a). Apo(a)
CC       fragments accumulate in atherosclerotic lesions, where they may
CC       promote thrombogenesis. O-glycosylation may limit the extent of
CC       proteolytic fragmentation.
CC   -!- SIMILARITY: CONTAINS 38 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06290; CAA29618.1; -.
DR   PIR; S00657; S00657.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.226; -.
DR   MIM; 152200; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 38.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 38.
```

```
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 38.
DR   PROSITE; PS50070; KRINGLE_2; 38.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW   Kringle; Repeat; Atherosclerosis; Signal; Polymorphism.
FT   SIGNAL       1     19
FT   CHAIN       20   4548       APOLIPOPROTEIN(A).
FT   DOMAIN      20    130       KRINGLE TYPE IV, 1.
FT   DOMAIN     131    244       KRINGLE TYPE IV, 2.
FT   DOMAIN     245    358       KRINGLE TYPE IV, 3.
FT   DOMAIN     359    472       KRINGLE TYPE IV, 4.
FT   DOMAIN     473    586       KRINGLE TYPE IV, 5.
FT   DOMAIN     587    700       KRINGLE TYPE IV, 6.
FT   DOMAIN     701    814       KRINGLE TYPE IV, 7.
FT   DOMAIN     815    928       KRINGLE TYPE IV, 8.
FT   DOMAIN     929   1042       KRINGLE TYPE IV, 9.
FT   DOMAIN    1043   1156       KRINGLE TYPE IV, 10.
FT   DOMAIN    1157   1270       KRINGLE TYPE IV, 11.
FT   DOMAIN    1271   1384       KRINGLE TYPE IV, 12.
FT   DOMAIN    1385   1498       KRINGLE TYPE IV, 13.
FT   DOMAIN    1499   1612       KRINGLE TYPE IV, 14.
FT   DOMAIN    1613   1726       KRINGLE TYPE IV, 15.
FT   DOMAIN    1727   1840       KRINGLE TYPE IV, 16.
FT   DOMAIN    1841   1954       KRINGLE TYPE IV, 17.
FT   DOMAIN    1955   2068       KRINGLE TYPE IV, 18.
FT   DOMAIN    2069   2182       KRINGLE TYPE IV, 19.
FT   DOMAIN    2183   2296       KRINGLE TYPE IV, 20.
FT   DOMAIN    2297   2410       KRINGLE TYPE IV, 21.
FT   DOMAIN    2411   2524       KRINGLE TYPE IV, 22.
FT   DOMAIN    2525   2638       KRINGLE TYPE IV, 23.
FT   DOMAIN    2639   2752       KRINGLE TYPE IV, 24.
FT   DOMAIN    2753   2866       KRINGLE TYPE IV, 25.
FT   DOMAIN    2867   2980       KRINGLE TYPE IV, 26.
FT   DOMAIN    2981   3094       KRINGLE TYPE IV, 27.
FT   DOMAIN    3095   3208       KRINGLE TYPE IV, 28.
FT   DOMAIN    3209   3322       KRINGLE TYPE IV, 29.
FT   DOMAIN    3323   3436       KRINGLE TYPE IV, 30.
FT   DOMAIN    3437   3550       KRINGLE TYPE IV, 31.
FT   DOMAIN    3551   3664       KRINGLE TYPE IV, 32.
FT   DOMAIN    3665   3770       KRINGLE TYPE IV, 33.
FT   DOMAIN    3771   3884       KRINGLE TYPE IV, 34.
FT   DOMAIN    3885   3998       KRINGLE TYPE IV, 35.
FT   DOMAIN    3999   4112       KRINGLE TYPE IV, 36.
FT   DOMAIN    4113   4226       KRINGLE TYPE IV, 37.
FT   DOMAIN    4227   4327       KRINGLE TYPE V.
FT   DOMAIN    4328   4548       SERINE PROTEASE.
FT   ACT_SITE  4369   4369       CHARGE RELAY SYSTEM.
FT   ACT_SITE  4412   4412       CHARGE RELAY SYSTEM.
FT   ACT_SITE  4498   4498       CHARGE RELAY SYSTEM.
FT   VARIANT   4193   4193       W -> R (LOSS OF LYSINE-SEPHAROSE
FT                               BINDING).
FT                               /FTId=VAR_006633.
SQ   SEQUENCE   4548 AA;  501313 MW;  96921BE96A465C5F CRC64;

  Query Match             84.0%;  Score 63;  DB 1;  Length 4548;
  Best Local Similarity   81.8%;  Pred. No. 0.063;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           |||||| ||:
Db    4279 RNPDGDINGPW 4289


RESULT 10
PLMN_ERIEU
ID   PLMN_ERIEU     STANDARD;      PRT;   810 AA.
AC   Q29485;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Erinaceus europaeus (Western European hedgehog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
OX   NCBI_TaxID=9365;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96025778; PubMed=7592597;
RA   Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA   Byrne C.D., Fong K.J., Meer K., Patthy L.;
RT   "The recurring evolution of lipoprotein(a). Insights from cloning of
RT   hedgehog apolipoprotein(a).";
RL   J. Biol. Chem. 270:24004-24009(1995).
RN   [2]
RP   REVISIONS.
RA   Lawn R.M.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U33171; AAC48717.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL       1     19       BY SIMILARITY.
FT   CHAIN       20    810       PLASMINOGEN.
FT   CHAIN       20    582       PLASMIN HEAVY CHAIN A (BY SIMILARITY).
FT   CHAIN      583    810       PLASMIN LIGHT CHAIN B (BY SIMILARITY).
FT   DOMAIN     583    810       SERINE PROTEASE.
FT   DOMAIN     103    181       KRINGLE 1.
```

```
FT   DOMAIN       185    262       KRINGLE 2.
FT   DOMAIN       275    352       KRINGLE 3.
FT   DOMAIN       379    456       KRINGLE 4.
FT   DOMAIN       482    561       KRINGLE 5.
FT   ACT_SITE     622    622       CHARGE RELAY SYSTEM.
FT   ACT_SITE     665    665       CHARGE RELAY SYSTEM.
FT   ACT_SITE     760    760       CHARGE RELAY SYSTEM.
FT   CARBOHYD     339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   810 AA;  90902 MW;  8E75780946017A16 CRC64;

  Query Match             80.0%;  Score 60;  DB 1;  Length 810;
  Best Local Similarity   81.8%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||  |||
Db     533 RNPDGDANGPW 543


RESULT 11
PLMN_PIG
ID   PLMN_PIG        STANDARD;      PRT;   790 AA.
AC   P06867;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE OF 1-560.
RA   Schaller J., Marti T., Roesselet S.J., Kaempfer U., Rickli E.E.;
RT   "Amino acid sequence of the heavy chain of porcine plasmin. Comparison
RT   of the carbohydrate attachment sites with the human and bovine
RT   species.";
RL   Fibrinolysis 1:91-102(1987).
RN   [2]
RP   SEQUENCE OF 450-790.
RX   MEDLINE=85203907; PubMed=3846533;
RA   Marti T., Schaller J., Rickli E.E.;
RT   "Determination of the complete amino-acid sequence of porcine
RT   miniplasminogen.";
RL   Eur. J. Biochem. 149:279-285(1985).
RN   [3]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=88185329; PubMed=3356193;
RA   Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA   Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT   "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT   plasminogen. Species specificity in relation to sialylation and
RT   fucosylation patterns.";
RL   Eur. J. Biochem. 173:57-63(1988).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLLACTOSAMINE, SIALIC ACID AND
CC       IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAC
CC       DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
CC       (MICROHETEROGENEITY).
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
DR   PIR; A25834; A25834.
DR   PIR; S03733; S03733.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   GlycoSuiteDB; P06867; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   CHAIN          1    560       PLASMIN HEAVY CHAIN A.
FT   CHAIN        561    790       PLASMIN LIGHT CHAIN B.
FT   DOMAIN       561    790       SERINE PROTEASE.
FT   DOMAIN        84    162       KRINGLE 1.
FT   DOMAIN       166    243       KRINGLE 2.
FT   DOMAIN       256    333       KRINGLE 3.
FT   DOMAIN       358    435       KRINGLE 4.
FT   DOMAIN       461    540       KRINGLE 5.
FT   ACT_SITE     602    602       CHARGE RELAY SYSTEM.
FT   ACT_SITE     645    645       CHARGE RELAY SYSTEM.
FT   ACT_SITE     740    740       CHARGE RELAY SYSTEM.
FT   CARBOHYD     289    289       N-LINKED (GLCNAC. . .).
FT                                 /FTId=CAR_000019.
FT   CARBOHYD     340    340       O-LINKED (GALNAC. . .).
FT                                 /FTId=CAR_000020.
SQ   SEQUENCE   790 AA;  88592 MW;  F04EA06E74BCD58E CRC64;

  Query Match             76.0%;  Score 57;  DB 1;  Length 790;
  Best Local Similarity   81.8%;  Pred. No. 0.1;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||||||  |||
Db     512 RNPDGDDNGPW 522


RESULT 12
HGFL_MOUSE
ID   HGFL_MOUSE      STANDARD;      PRT;   716 AA.
AC   P26928;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor-like protein precursor (Macrophage
DE   stimulatory protein) (MSP).
GN   MST1 OR HGFL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Liver;
RX   MEDLINE=92002017; PubMed=1832957;
RA   Friezner Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT   "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT   growth factor-like protein: expression during development.";
```

```
RL   Biochemistry 30:9781-9791(1991).
CC   -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC       CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC       CONSERVED.
CC   -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC       ADRENAL.
CC   -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC       JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC       STABLE AFTERWARDS.
CC   -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC       HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC       POLYPEPTIDES.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74180; AAA50166.1; -.
DR   EMBL; M74181; AAA50167.1; -.
DR   HSSP; P00747; 1KRN.
DR   MEROPS; S01.975; -.
DR   MGD; MGI:96080; Hgfl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Kringle; Glycoprotein; Serine protease homolog; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    716       HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
FT   DOMAIN      19    109       PAP.
FT   DOMAIN     110    186       KRINGLE 1.
FT   DOMAIN     191    268       KRINGLE 2.
FT   DOMAIN     292    370       KRINGLE 3.
FT   DOMAIN     379    457       KRINGLE 4.
FT   DOMAIN     489    716       SERINE PROTEASE-LIKE.
FT   DISULFID    56     78       BY SIMILARITY.
FT   DISULFID    60     66       BY SIMILARITY.
FT   DISULFID   110    186       BY SIMILARITY.
FT   DISULFID   131    169       BY SIMILARITY.
FT   DISULFID   157    181       BY SIMILARITY.
FT   DISULFID   191    268       BY SIMILARITY.
FT   DISULFID   194    333       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   212    251       BY SIMILARITY.
FT   DISULFID   240    263       BY SIMILARITY.
FT   DISULFID   292    370       BY SIMILARITY.
FT   DISULFID   313    352       BY SIMILARITY.
FT   DISULFID   341    364       BY SIMILARITY.
FT   DISULFID   379    457       BY SIMILARITY.
FT   DISULFID   400    440       BY SIMILARITY.
FT   DISULFID   428    452       BY SIMILARITY.
FT   DISULFID   477    593       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   512    528       BY SIMILARITY.
FT   DISULFID   607    672       BY SIMILARITY.
FT   DISULFID   637    651       BY SIMILARITY.
FT   DISULFID   662    690       BY SIMILARITY.
FT   CARBOHYD    72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   305    305       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   620    620       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    19     19       P -> Q (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   716 AA;  80588 MW;  BBCE02EF85213ACC CRC64;

  Query Match             74.7%;  Score 56;  DB 1;  Length 716;
  Best Local Similarity   81.8%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           |||||| |||
Db     429 RNPDGDSHGPW 439


RESULT 13
THRB_BOVIN
ID   THRB_BOVIN     STANDARD;      PRT;   625 AA.
AC   P00735;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88245190; PubMed=3379642;
RA   Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
RT   "Structure and evolution of the bovine prothrombin gene.";
RL   J. Mol. Biol. 200:31-45(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84203525; PubMed=6326805;
RA   McGillivray R.T.A., Davie E.W.;
RT   "Characterization of bovine prothrombin mRNA and its translation
RT   product.";
RL   Biochemistry 23:1626-1634(1984).
RN   [3]
RP   SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RA   Magnusson S., Sottrup-Jensen L., Petersen T.E., Claeys H.;
RL   (In) Hemker H.C., Veltkamp J.J. (eds.);
RL   Boerhaave symposium on prothrombin and related coagulation factors,
RL   pp.25-46, Leiden University Press, Leiden (1975).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=86296631; PubMed=3741841;
RA   Park C.H., Tulinsky A.;
RT   "Three-dimensional structure of the kringle sequence: structure of
RT   prothrombin fragment 1.";
RL   Biochemistry 25:3977-3982(1986).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=91311686; PubMed=1856869;
RA   Seshadri T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;
RT   "Structure of bovine prothrombin fragment 1 refined at 2.25-A
RT   resolution.";
RL   J. Mol. Biol. 220:481-494(1991).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=92190185; PubMed=1547238;
RA   Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;
RT   "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-
RT   prothrombin fragment 1.";
RL   Biochemistry 31:2554-2566(1992).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
```

```
RX   MEDLINE=92218459; PubMed=1560020;
RA   Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
RT   "The structure of residues 7-16 of the A alpha-chain of human
RT   fibrinogen bound to bovine thrombin at 2.3-A resolution.";
RL   J. Biol. Chem. 267:7911-7920(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=92389319; PubMed=1518046;
RA   Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
RA   Martin P.D., Edwards B.F.P., Bode W.;
RT   "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes
RT   formed with the benzamidine and arginine-based thrombin inhibitors
RT   NAPAP, 4-TAPAP and MQPA. A starting point for improving
RT   antithrombotics.";
RL   J. Mol. Biol. 226:1085-1089(1992).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
RX   MEDLINE=97102783; PubMed=8947023;
RA   van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
RA   Hoffken W., Huber R.;
RT   "The ornithodorin-thrombin crystal structure, a key to the TAP
RT   enigma?";
RL   EMBO J. 15:6011-6017(1996).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
RX   MEDLINE=98004486; PubMed=9342325;
RA   Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
RA   Huber R., Bode W.;
RT   "Structure of the thrombin complex with triabin, a lipocalin-like
RT   exosite-binding inhibitor derived from a triatomine bug.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
RN   [11]
RP   GENE STRUCTURE.
RX   MEDLINE=86077733; PubMed=3000440;
RA   Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;
RT   "Characterization of the bovine prothrombin gene.";
RL   Biochemistry 24:6854-6861(1985).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   -!- DATABASE: NAME=ProZyme technical fact sheet;
CC       WWW="http://www.prozyme.com/technical/thrombindata.html".
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V00135; CAA23451.1; -.
DR   EMBL; J00041; AAA30781.1; -.
DR   PIR; A00915; TBBO.
DR   PIR; S02537; S02537.
DR   PDB; 1BBR; 31-JAN-94.
DR   PDB; 1ETR; 31-JAN-94.
DR   PDB; 1ETS; 31-JAN-94.
DR   PDB; 1ETT; 31-JAN-94.
DR   PDB; 1HRT; 31-JAN-94.
DR   PDB; 2PF1; 31-JAN-94.
DR   PDB; 2PF2; 31-JAN-94.
DR   PDB; 2SPT; 31-MAY-94.
DR   PDB; 1MKW; 07-JUL-97.
DR   PDB; 1MKX; 07-JUL-97.
DR   PDB; 1TBQ; 14-OCT-96.
DR   PDB; 1TBR; 14-OCT-96.
DR   PDB; 1TOC; 23-JUL-97.
DR   PDB; 1VIT; 21-APR-97.
DR   PDB; 1YCP; 06-MAY-98.
DR   PDB; 1A0H; 17-JUN-98.
DR   PDB; 1AVG; 16-FEB-99.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal; 3D-structure.
FT   SIGNAL       1    24       POTENTIAL.
FT   PROPEP      25    43
FT   CHAIN       44   625       PROTHROMBIN.
FT   PEPTIDE     44   199       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE    200   317       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN      318   366       THROMBIN LIGHT CHAIN (A).
FT   CHAIN      367   625       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN     109   187       KRINGLE 1.
FT   DOMAIN     214   292       KRINGLE 2.
FT   DOMAIN     367   625       SERINE PROTEASE.
FT   SITE       199   200       CLEAVAGE (BY THROMBIN).
FT   SITE       317   318       CLEAVAGE (BY FACTOR XA).
FT   SITE       366   367       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE   409   409       CHARGE RELAY SYSTEM.
FT   ACT_SITE   465   465       CHARGE RELAY SYSTEM.
FT   ACT_SITE   571   571       CHARGE RELAY SYSTEM.
FT   MOD_RES     50    50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     51    51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     58    58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60    60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     63    63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     64    64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     69    69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     70    70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     73    73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     76    76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD   120   120       N-LINKED (GLCNAC. . .).
```

```
  Query Match             73.3%;  Score 55;  DB 1;  Length 625;
  Best Local Similarity   72.7%;  Pred. No. 0.17;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPW 11
          ||||| : |||
Db    159 RNPDGSITGPW 169


RESULT 14
HGF_HUMAN
ID   HGF_HUMAN      STANDARD;      PRT;   728 AA.
AC   P14210; Q9UDU6; Q9BYL9;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor precursor (Scatter factor) (SF)
DE   (Hepatopoeitin A).
GN   HGF OR HPTA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91340155; PubMed=1831432;
RA   Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT   "Organization of the human hepatocyte growth factor-encoding gene.";
RL   Gene 102:213-219(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=89392017; PubMed=2528952;
RA   Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA   Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RA   Gohda E., Daikuhara Y., Kitamura N.;
RT   "Molecular cloning and sequence analysis of cDNA for human hepatocyte
RT   growth factor.";
RL   Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=91025062; PubMed=2145836;
RA   Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA   Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT   "Isolation and expression of cDNA for different forms of hepatocyte
RT   growth factor from human leukocyte.";
RL   Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN   [4]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RC   TISSUE=Liver;
RX   MEDLINE=90066676; PubMed=2531289;
RA   Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA   Sugimura A., Tashiro K., Shimizu S.;
RT   "Molecular cloning and expression of human hepatocyte growth factor.";
RL   Nature 342:440-443(1989).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryonic fibroblast;
RX   MEDLINE=91334393; PubMed=1831266;
RA   Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA   Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA   Birchmeier W.;
RT   "Evidence for the identity of human scatter factor and human
RT   hepatocyte growth factor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Courtney L., Elliot G., Angell S.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 249-695 FROM N.A.
RX   MEDLINE=91369928; PubMed=1832556;
RA   Miyazawa K., Kitamura A., Kitamura N.;
RT   "Structural organization and the transcription initiation site of the
RT   human hepatocyte growth factor gene.";
RL   Biochemistry 30:9170-9176(1991).
RN   [8]
RP   SIGNAL SEQUENCE CLEAVAGE SITE.
RX   MEDLINE=91207365; PubMed=1826837;
RA   Yoshiyama Y., Arakakai N., Naka D., Takahashi K., Hirono S., Kondo J.,
RA   Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
RA   Hishida T., Daikuhara Y.;
RT   "Identification of the N-terminal residue of the heavy chain of both
RT   native and recombinant human hepatocyte growth factor.";
RL   Biochem. Biophys. Res. Commun. 175:660-667(1991).
RN   [9]
RP   CARBOHYDRATE-LINKAGE SITE 476.
RX   MEDLINE=93129192; PubMed=1482348;
RA   Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
RA   Nakamura T., Shimizu S.;
RT   "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
RT   on the alpha chain.";
RL   Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
RN   [10]
RP   MUTAGENESIS.
RX   MEDLINE=92331602; PubMed=1321034;
RA   Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
RA   Baker J.B., Godowski P.J.;
RT   "Structure-function analysis of hepatocyte growth factor:
RT   identification of variants that lack mitogenic activity yet retain
RT   high affinity receptor binding.";
RL   EMBO J. 11:2503-2510(1992).
RN   [11]
RP   STRUCTURE BY NMR OF 31-127.
RX   MEDLINE=98154323; PubMed=9493272;
RA   Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
RA   Rubin J.S., Bottaro D.P., Byrd R.A.;
RT   "The solution structure of the N-terminal domain of hepatocyte growth
RT   factor reveals a potential heparin-binding site.";
RL   Structure 6:109-116(1998).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
RX   MEDLINE=99036858; PubMed=9817840;
RA   Ultsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
RT   "Crystal structure of the NK1 fragment of human hepatocyte growth
RT   factor at 2.0-A resolution.";
RL   Structure 6:1383-1393(1998).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC       AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC       IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90334; BAA14348.1; -.
DR   EMBL; D90318; BAA14348.1; JOINED.
DR   EMBL; D90319; BAA14348.1; JOINED.
DR   EMBL; D90320; BAA14348.1; JOINED.
DR   EMBL; D90322; BAA14348.1; JOINED.
DR   EMBL; D90323; BAA14348.1; JOINED.
DR   EMBL; D90324; BAA14348.1; JOINED.
DR   EMBL; D90325; BAA14348.1; JOINED.
DR   EMBL; D90326; BAA14348.1; JOINED.
```

```
DR   EMBL; D90327; BAA14348.1; JOINED.
DR   EMBL; D90328; BAA14348.1; JOINED.
DR   EMBL; D90329; BAA14348.1; JOINED.
DR   EMBL; D90330; BAA14348.1; JOINED.
DR   EMBL; D90331; BAA14348.1; JOINED.
DR   EMBL; D90332; BAA14348.1; JOINED.
DR   EMBL; D90333; BAA14348.1; JOINED.
DR   EMBL; M29145; AAA52650.1; -.
DR   EMBL; M60718; AAA52648.1; -.
DR   EMBL; X16323; CAA34387.1; -.
DR   EMBL; M73239; AAA64239.1; -.
DR   EMBL; M73240; AAA64297.1; -.
DR   EMBL; AC004960; AAC71655.1; -.
DR   EMBL; M75983; AAG53460.1; -.
DR   EMBL; M75972; AAG53460.1; JOINED.
DR   EMBL; M75973; AAG53460.1; JOINED.
DR   EMBL; M75974; AAG53460.1; JOINED.
DR   EMBL; M75975; AAG53460.1; JOINED.
DR   EMBL; M75976; AAG53460.1; JOINED.
DR   EMBL; M75977; AAG53460.1; JOINED.
DR   EMBL; M75978; AAG53460.1; JOINED.
DR   EMBL; M75979; AAG53460.1; JOINED.
DR   EMBL; M75980; AAG53460.1; JOINED.
DR   EMBL; M75981; AAG53460.1; JOINED.
DR   EMBL; M75982; AAG53460.1; JOINED.
DR   PIR; JH0579; JH0579.
DR   PIR; S06794; S06794.
DR   PDB; 2HGF; 24-JUN-98.
DR   PDB; 1BHT; 18-NOV-98.
DR   MEROPS; S01.976; -.
DR   GlycoSuiteDB; P14210; -.
DR   MIM; 142409; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW   Signal; 3D-structure.
FT   SIGNAL       1     31
FT   CHAIN       32    494       HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT   CHAIN      495    728       HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT   MOD_RES     32     32       PYRROLIDONE CARBOXYLIC ACID.
FT   DOMAIN      32    127       PAP.
FT   DOMAIN     128    206       KRINGLE 1.
FT   DOMAIN     211    288       KRINGLE 2.
FT   DOMAIN     305    383       KRINGLE 3.
FT   DOMAIN     391    469       KRINGLE 4.
FT   DOMAIN     495    728       SERINE PROTEASE-LIKE.
FT   DISULFID    70     96
FT   DISULFID    74     84
FT   DISULFID   128    206
FT   DISULFID   149    189
FT   DISULFID   177    201
FT   DISULFID   487    604       INTERCHAIN (BY SIMILARITY).

  Query Match             69.3%;  Score 52;  DB 1;  Length 728;
  Best Local Similarity   72.7%;  Pred. No. 0.61;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||

Db     178 RNPRGEEGGPW 188


RESULT 15
HGF_MOUSE
ID   HGF_MOUSE      STANDARD;      PRT;   728 AA.
AC   Q08048; Q64007; Q61662;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor precursor (Scatter factor) (SF)
DE   (Hepatopoeitin A).
GN   HGF.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 496-504.
RC   TISSUE=Mammary fibroblast;
RX   MEDLINE=94183257; PubMed=8135822;
RA   Sasaki M., Nishio M., Sasaki T., Enami J.;
RT   "Identification of mouse mammary fibroblast-derived mammary growth
RT   factor as hepatocyte growth factor.";
RL   Biochem. Biophys. Res. Commun. 199:772-779(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=94363381; PubMed=8081873;
RA   Lee C.C., Kozak C.A., Yamada K.M.;
RT   "Structure, genetic mapping, and expression of the mouse Hgf/scatter
RT   factor gene.";
RL   Cell Adhes. Commun. 1:101-111(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=94060105; PubMed=8241272;
RA   Liu Y., Michalopoulos G.K., Zarnegar R.;
RT   "Molecular cloning and characterization of cDNA encoding mouse
RT   hepatocyte growth factor.";
RL   Biochim. Biophys. Acta 1216:299-303(1993).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC       AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC       IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY
CC       ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE
CC       LONG FORM.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10212; BAA01064.1; -.
DR   EMBL; D10213; BAA01065.1; -.
DR   EMBL; S71816; AAB31855.1; -.
DR   EMBL; X72307; CAA51054.1; ALT_INIT.
DR   HSSP; P14210; 1BHT.
DR   MGD; MGI:96079; Hgf.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
```

```
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW   Signal; Alternative splicing.
FT   SIGNAL        1     32       BY SIMILARITY.
FT   CHAIN        33    495       HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT   CHAIN       496    728       HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT   MOD_RES      33     33       PYRROLIDONE CARBOXYLIC ACID
FT                                (BY SIMILARITY).
FT   DOMAIN       33    128       PAP.
FT   DOMAIN      129    207       KRINGLE 1.
FT   DOMAIN      212    289       KRINGLE 2.
FT   DOMAIN      306    384       KRINGLE 3.
FT   DOMAIN      392    470       KRINGLE 4.
FT   DOMAIN      496    728       SERINE PROTEASE-LIKE.
FT   DISULFID     71     97       BY SIMILARITY.
FT   DISULFID     75     85       BY SIMILARITY.
FT   DISULFID    488    607       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    569    569       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    656    656       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    163    167       MISSING (IN SHORT ISOFORM).
FT   CONFLICT    344    344       N -> K (IN REF. 2).
FT   CONFLICT    479    479       V -> L (IN REF. 2).
FT   CONFLICT    564    564       R -> H (IN REF. 3).
SQ   SEQUENCE   728 AA;  82944 MW;  A0381FC497534328 CRC64;

  Query Match             69.3%;  Score 52;  DB 1;  Length 728;
  Best Local Similarity   72.7%;  Pred. No. 0.61;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPW 11
           ||| |: ||||
Db     179 RNPRGEEGGPW 189


Search completed: November  8, 2002, 09:33:52
Job time : 10.3333 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:31:21 ; Search time 31 Seconds
(without alignments)
82.410 Million cell updates/sec

Title: US-09-657-431-9
Perfect score: 138
Sequence: 1 RNPDGDVGGPWAYTTNPRKLYDY 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	23	22	AAB92096	Laminin fragment S
2	138	100.0	23	22	AAB36570	Mammalian kringle
3	138	100.0	24	22	AAB92090	Laminin fragment S
4	138	100.0	24	22	AAB92095	Laminin fragment S
5	138	100.0	24	22	AAB36565	Mammalian kringle
6	134	97.1	79	18	AAW19256	Human plasminogen
7	134	97.1	90	21	AAB01914	Human plasminogen
8	134	97.1	91	21	AAY58868	Human plasminogen
9	134	97.1	93	21	AAB01917	Human plasminogen
10	134	97.1	95	21	AAB01913	Human plasminogen
11	134	97.1	98	21	AAB01916	Human plasminogen
12	134	97.1	101	18	AAW34286	Human kringle 5 pe
13	134	97.1	101	21	AAB01890	Human plasminogen
14	134	97.1	101	21	AAB01912	Human plasminogen
15	134	97.1	104	21	AAB01915	Human plasminogen
16	134	97.1	189	21	AAB01918	Human plasminogen
17	134	97.1	192	21	AAB01919	Human plasminogen
18	134	97.1	266	22	AAU32126	Novel human secret
19	134	97.1	266	22	AAU32129	Novel human secret
20	134	97.1	266	22	AAU32136	Novel human secret
21	134	97.1	271	21	AAB08407	A human angiogenes
22	134	97.1	357	20	AAY25408	Human tissue facto
23	134	97.1	371	13	AAR22502	[GARSYQ]-[Plasmino
24	134	97.1	380	13	AAR22504	[GARSYQ]-[Plasmino
25	134	97.1	437	19	AAW51457	Human plasminogen
26	134	97.1	467	13	AAR22499	[GARSYQ]-[Plasmino
27	134	97.1	476	13	AAR22503	[GARSYQ]-[Plasmino
28	134	97.1	566	20	AAY02100	A multifunctional
29	134	97.1	790	15	AAR60519	Human 'Glu' plasmi
30	134	97.1	790	22	AAB36562	Mammalian kringle
31	134	97.1	791	21	AAB01887	Human plasminogen,
32	134	97.1	791	21	AAY99589	Human plasminogen.
33	134	97.1	791	21	AAY50867	Human plasminogen
34	134	97.1	791	22	AAG67223	Amino acid sequenc
35	134	97.1	807	13	AAR20013	PA mutant Plg 1-54
36	134	97.1	810	11	AAR08065	Human plasminogen
37	134	97.1	810	12	AAR13219	R561G human plasmi
38	134	97.1	810	12	AAR12406	R561E human plasmi
39	134	97.1	810	12	AAR13220	R561S human plasmi
40	134	97.1	810	12	AAR13221	Human plasminogen
41	134	97.1	810	12	AAR12938	Plasminogen mutein
42	134	97.1	810	14	AAR34428	Sequence encoded b
43	134	97.1	810	18	AAW31169	Plasminogen protei
44	134	97.1	810	20	AAY08685	Human plasminogen
45	134	97.1	810	20	AAY02114	SEQ ID 77 of WO991
46	134	97.1	810	21	AAY82690	Human plasminogen
47	134	97.1	810	21	AAY53867	Amino acid sequenc
48	134	97.1	811	12	AAR12933	Plasminogen mutein
49	134	97.1	811	12	AAR12939	Plasminogen mutein
50	134	97.1	811	12	AAR12943	Plasminogen mutein

ALIGNMENTS

RESULT 1
AAB92096
ID AAB92096 standard; Peptide; 23 AA.
XX
AC AAB92096;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1272.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 611; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 23 AA;

Query Match 100.0%; Score 138; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||||||||||||||||
Db 1 RNPDGDVGGPWAYTTNPRKLYDY 23

RESULT 2
AAB36570
ID AAB36570 standard; Peptide; 23 AA.
XX
AC AAB36570;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:9.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 5; Page 9; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 23 AA;

Query Match 100.0%; Score 138; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||||||||||||||||
Db 1 RNPDGDVGGPWAYTTNPRKLYDY 23

RESULT 3
AAB92090
ID AAB92090 standard; Peptide; 24 AA.
XX
AC AAB92090;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1266.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX

DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 609; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 138; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||||||||||||||||
Db 1 RNPDGDVGGPWAYTTNPRKLYDY 23

RESULT 4
AAB92095
ID AAB92095 standard; Peptide; 24 AA.
XX
AC AAB92095;
XX
DT 22-JUN-2001 (first entry)
XX
DE Laminin fragment SEQ ID NO:1271.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 611; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 138; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||||||||||||||||
Db 1 RNPDGDVGGPWAYTTNPRKLYDY 23

RESULT 5
AAB36565
ID AAB36565 standard; Peptide; 24 AA.
XX
AC AAB36565;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 peptide SEQ ID NO:4.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Claim 5; Page 9; 82pp; English.
XX

CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 138; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||||||||||||||||
Db 1 RNPDGDVGGPWAYTTNPRKLYDY 23

RESULT 6
AAW19256
ID AAW19256 standard; Peptide; 79 AA.
XX
AC AAW19256;
XX
DT 27-FEB-1998 (first entry)
XX
DE Human plasminogen Kringle 5 fragment.
XX
KW Plasminogen; Kringle 5; cell proliferation inhibitor; angiogenesis;
KW diagnosis; therapeutic.
XX
OS Homo sapiens.
XX
PN WO9723500-A1.
XX
PD 03-JUL-1997.
XX
PF 13-DEC-1996; 96WO-US20447.
XX
PR 12-DEC-1996; 96US-0763528.
PR 13-DEC-1995; 95US-0008519.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Cao Y, Folkman MJ;
XX
DR WPI; 1997-350965/32.
XX
PT Plasminogen Kringle 5 peptide - which inhibits endothelial cell
PT proliferation, useful to treat angiogenesis mediated diseases and in
PT detection and diagnosis
XX
PS Claim 1; Page 8; 51pp; English.
XX
CC This sequence is an isolated fragment of the Kringle 5 peptide
CC corresponding to amino acid 462 of the human plasminogen protein which
CC can be used in a novel method to inhibit endothelial cell proliferation
CC activity. The protein can be used to treat angiogenesis mediated
CC diseases, e.g. haemangioma, solid tumours, leukaemia, metastasis,
CC telangiectasia, psoriasis, scleroderma, pyogenic granuloma, myocardial
CC angiogenesis, plaque neovascularisation, coronary or cerebral
CC collaterals, arteriovenous malformations, ischemic limb angiogenesis,
CC corneal diseases, rubeosis, neovascular glaucoma, diabetic retinopathy,
CC retrolental fibroplasia, arthritis, diabetic neovascularisation,
CC muscular degeneration, peptic ulcer, Helicobacter related disease,
CC fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
CC menstruation, placentation or cat scratch fever, and to stimulate wound
CC healing. The protein and antibodies generated from it can be used to
CC screen for agonists and antagonists or in detection, imaging and
CC diagnosis.
XX
SQ Sequence 79 AA;

Query Match 97.1%; Score 134; DB 18; Length 79;
Best Local Similarity 95.7%; Pred. No. 4.8e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 52 RNPDGDVGGPWCYTTNPRKLYDY 74

RESULT 7
AAB01914
ID AAB01914 standard; Protein; 90 AA.
XX
AC AAB01914;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5 (Val454-Ala543).
XX
KW Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 17; Page -; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours

```
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   90 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 90;
  Best Local Similarity   95.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      60 RNPDGDVGGPWCYTTNPRKLYDY 82


RESULT 8
AAY58868
ID   AAY58868 standard; Protein; 91 AA.
XX
AC   AAY58868;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Human plasminogen mature polypeptide.
XX
KW   Anti-angiogenic; angiogenesis inhibitor; cancer; tumour; therapy;
KW   plasminogen; human.
XX
OS   Homo sapiens.
XX
PN   WO200004052-A2.
XX
PD   27-JAN-2000.
XX
PF   16-JUL-1999;   99WO-GB02292.
XX
PR   16-JUL-1998;   98GB-0015505.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Smith RAG,  Bright JR,  Steward M,  Cox VF;
XX
DR   WPI; 2000-182406/16.
XX
PT   New soluble derivative of anti-angiogenic polypeptide useful for
PT   treatment of primary or secondary cancers, contains covalently attached
PT   membrane-binding elements for targeting  -
XX
PS   Example 3; Page 21; 36pp; English.
XX
CC   The present sequence is that of human plasminogen mature
CC   polypeptide.  The invention relates to new soluble derivatives (I)
CC   of anti-angiogenic polypeptides.  (I) comprise 2 or more
CC   heterologous membrane binding elements (MBEs, see AAY58855-61) with
CC   low membrane affinity that are covalently attached to a soluble
CC   anti-angiogenic polypeptide such as a non-catalytic region of human
CC   plasminogen.  The MBEs interact independently with thermodynamic
CC   additivity, with components of the vascular endothelium.  (I)
CC   provide targeted delivery of the anti-angiogenic polypeptide to
CC   cell membranes and sites of active angiogenesis, particularly the
CC   vascular endothelium, and therefore increase the local
CC   concentration and reduce the risk of adverse effects on normal
CC   processes elsewhere in the vasculature.  They are used in a claimed
CC   method of treatment of primary or secondary tumour.
XX
SQ   Sequence   91 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 91;
  Best Local Similarity   95.7%;  Pred. No. 5.6e-10;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      56 RNPDGDVGGPWCYTTNPRKLYDY 78


RESULT 9
AAB01917
ID   AAB01917 standard; Protein; 93 AA.
XX
AC   AAB01917;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 (Val454-Phe546).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
```

CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
SQ Sequence 93 AA;

Query Match 97.1%; Score 134; DB 21; Length 93;
Best Local Similarity 95.7%; Pred. No. 5.7e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      60 RNPDGDVGGPWCYTTNPRKLYDY 82
```

RESULT 10
AAB01913
ID AAB01913 standard; Protein; 95 AA.
XX
AC AAB01913;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5 (Val449-Ala543).
XX
KW Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 17; Page -; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC Syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived

CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
SQ Sequence 95 AA;

Query Match 97.1%; Score 134; DB 21; Length 95;
Best Local Similarity 95.7%; Pred. No. 5.8e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      65 RNPDGDVGGPWCYTTNPRKLYDY 87
```

RESULT 11
AAB01916
ID AAB01916 standard; Protein; 98 AA.
XX
AC AAB01916;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5 (Val449-Phe546).
XX
KW Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 17; Page -; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC Syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived

CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
SQ Sequence 98 AA;

Query Match 97.1%; Score 134; DB 21; Length 98;
Best Local Similarity 95.7%; Pred. No. 6e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 65 RNPDGDVGGPWCYTTNPRKLYDY 87

RESULT 12
AAW34286
ID AAW34286 standard; protein; 101 AA.
XX
AC AAW34286;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human kringle 5 peptide fragment.
XX
KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
KW autoimmune disease; ocular disease; capillary proliferation; therapy;
KW kringle 5 receptor.
XX
OS Homo sapiens.
XX
PN WO9741824-A2.
XX
PD 13-NOV-1997.
XX
PF 05-MAY-1997; 97WO-US07700.
XX
PR 03-APR-1997; 97US-0832087.
PR 03-MAY-1996; 96US-0643219.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ, Gubbins EJ, Wang J;
XX
DR WPI; 1997-558670/51.
XX
PT New kringle 5 peptide(s) and fusion proteins derived from
PT plasminogen - useful as anti-angiogenesis agents for treating
PT cancer, psoriasis, arthritis etc., including gene therapy
XX
PS Claim 35; Fig 2; 78pp; English.
XX
CC This sequence is a kringle 5 (K5) peptide homologous to human
CC plasminogen. K5 peptide fragments homologous to this sequence, are
CC anti-angiogenesis agents, specifically for treating or preventing cancer,
CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,
CC lymphomas, haemangiomas. They can also be used for treating or preventing
CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The
CC fragments can also be used to treat autoimmune or ocular diseases,
CC capillary proliferation within atherosclerotic plaque, haemophiliac
CC joints, wound granulation, ulcers etc., also as contraceptives that
CC inhibit ovulation and establishment of the placenta. K5 antisera or
CC (ant)agonists can be used similarly, optionally coupled to cytotoxic
CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound
CC healing. The K5 peptides are also used to raise specific antibodies (Ab),
CC for diagnosis and for affinity purification of K5 receptors. The K5
CC receptors may then be expressed in tumour cells to increase their
CC response to the peptides or used for identification of smaller
CC antagonists. The Ab are used to detect/quantify the peptides in
CC biological samples. The K5 peptides (and K5 fusion proteins) selectively
CC inhibit proliferation of endothelial cells with low toxicity against
CC normal cells. Typically they have 800-times greater inhibitory activity
CC against bovine capillary cells in vitro than kringle 1-4 peptides.
XX
SQ Sequence 101 AA;

Query Match 97.1%; Score 134; DB 18; Length 101;
Best Local Similarity 95.7%; Pred. No. 6.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 71 RNPDGDVGGPWCYTTNPRKLYDY 93

RESULT 13
AAB01890
ID AAB01890 standard; Protein; 101 AA.
XX
AC AAB01890;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 5, SEQ ID NO:34.
XX
KW Plasminogen; kringle 5 domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Disclosure; Fig 2; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC Syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01890-B01894 represent, respectively, human, murine, Rhesus
CC monkey, bovine and porcine plasminogen kringle 5 domains.
XX
SQ Sequence 101 AA;

```
  Query Match             97.1%;  Score 134;  DB 21;  Length 101;
  Best Local Similarity   95.7%;  Pred. No. 6.2e-10;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      71 RNPDGDVGGPWCYTTNPRKLYDY 93


RESULT 14
AAB01912
ID   AAB01912 standard; Protein; 101 AA.
XX
AC   AAB01912;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 (Val443-Ala543).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   101 AA;
```

```
  Query Match             97.1%;  Score 134;  DB 21;  Length 101;
  Best Local Similarity   95.7%;  Pred. No. 6.2e-10;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      71 RNPDGDVGGPWCYTTNPRKLYDY 93


RESULT 15
AAB01915
ID   AAB01915 standard; Protein; 104 AA.
XX
AC   AAB01915;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringle 5 (Val443-Phe546).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   104 AA;
```

```
  Query Match             97.1%;  Score 134;  DB 21;  Length 104;
  Best Local Similarity   95.7%;  Pred. No. 6.4e-10;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      71 RNPDGDVGGPWCYTTNPRKLYDY 93


RESULT 16
AAB01918
ID   AAB01918 standard; Protein; 189 AA.
XX
AC   AAB01918;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringles 4-5 (Val355-Ala543).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   189 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 189;
  Best Local Similarity   95.7%;  Pred. No. 1.2e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     159 RNPDGDVGGPWCYTTNPRKLYDY 181


RESULT 17
AAB01919
ID   AAB01919 standard; Protein; 192 AA.
XX
AC   AAB01919;
XX
DT   18-SEP-2000  (first entry)
XX
DE   Human plasminogen kringles 4-5 (Val355-Phe546).
XX
KW   Plasminogen; human; kringle domain; endothelial cell proliferation;
KW   angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Page -; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful
CC   for treating angiogenic diseases, primary and metastatic solid tumours
CC   and carcinomas of various organs such as breast, genital tract,
CC   endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC   arising from haematopoietic malignancies such as leukaemias and
CC   lymphomas. They are also used for the prophylaxis of various autoimmune
CC   diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC   (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC   Syndrome), diseases caused by excessive or abnormal stimulation of
CC   endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC   which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC   disease and ulcers). The peptides are also useful as a birth control
CC   agent which inhibits ovulation and establishment of the placenta.
CC   Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC   in an exemplification of the invention.
CC   Note: This sequence is not shown in the specification, but is derived
CC   from  the full length human plasminogen sequence (AAB01887) shown in
CC   figure 1.
XX
SQ   Sequence   192 AA;
```

```
  Query Match             97.1%;  Score 134;  DB 21;  Length 192;
  Best Local Similarity   95.7%;  Pred. No. 1.2e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     159 RNPDGDVGGPWCYTTNPRKLYDY 181


RESULT 18
AAU32126
ID   AAU32126 standard; Protein; 266 AA.
XX
AC   AAU32126;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #2617.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 558; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   266 AA;

  Query Match             97.1%;  Score 134;  DB 22;  Length 266;
  Best Local Similarity   95.7%;  Pred. No. 1.6e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     109 RNPDGDVGGPWCYTTNPRKLYDY 131


RESULT 19
AAU32129
ID   AAU32129 standard; Protein; 266 AA.
XX
AC   AAU32129;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #2620.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 559; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   266 AA;

  Query Match             97.1%;  Score 134;  DB 22;  Length 266;
  Best Local Similarity   95.7%;  Pred. No. 1.6e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     109 RNPDGDVGGPWCYTTNPRKLYDY 131


RESULT 20
AAU32136
ID   AAU32136 standard; Protein; 266 AA.
XX
AC   AAU32136;
XX
```

DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #2627.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 561; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 266 AA;

Query Match 97.1%; Score 134; DB 22; Length 266;
Best Local Similarity 95.7%; Pred. No. 1.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 109 RNPDGDVGGPWCYTTNPRKLYDY 131

RESULT 21
AAB08407
ID AAB08407 standard; Protein; 271 AA.
XX
AC AAB08407;
XX
DT 20-DEC-2000 (first entry)
XX
DE A human angiogenesis inhibitor designated KED.
XX
KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW adult respiratory distress syndrome; Castleman's disease; psoriasis;
KW hepatitis; aneurysm; renal disease; haemangioma.
XX
OS Homo sapiens.
XX
PN WO200049871-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-US04798.
XX
PR 24-FEB-1999; 99US-0121341.
PR 25-FEB-1999; 99US-0121633.
PR 18-NOV-1999; 99US-0166176.
XX
PA (FORD-) FORD HEALTH SYSTEM HENRY.
XX
PI Dou D, Chopp M, Wang L, Mikkelsen T;
XX
DR WPI; 2000-572016/53.
DR N-PSDB; AAA64013.
XX
PT Use of kringle protein and kringle derived from plasminogen and
PT composition comprising kringle proteins for treating tumor and
PT atherosclerosis, arthritis and retinopathy -
XX
PS Disclosure; Page 151; 163pp; English.
XX
CC The present sequence represents a human polypeptide which is a potent
CC angiogenesis inhibitor, designated KED. The KED polypeptide, kringle
CC proteins, or a kringle derived from human tissue plasminogen activator
CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC or tPA are useful for treating tumours, as well as atherosclerosis,
CC arthritis, retinopathy and other similar diseases. KED is also useful
CC for the treatment of diseases such as bronchial vascular congestion,
CC inflammatory bowel disease, adult respiratory distress syndrome,
CC Castleman's disease, psoriasis, hepatitis, aneurysm, renal disease
CC and haemangioma.
XX
SQ Sequence 271 AA;

Query Match 97.1%; Score 134; DB 21; Length 271;
Best Local Similarity 95.7%; Pred. No. 1.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 56 RNPDGDVGGPWCYTTNPRKLYDY 78

RESULT 22
AAY25408
ID AAY25408 standard; Protein; 357 AA.
XX
AC AAY25408;
XX
DT 08-SEP-1999 (first entry)
XX
DE Human tissue factor clone NuV129 protein.
XX
KW Tissue factor; human; thrombogenic; substructure; thrombose; tumour;
KW vasculative malformation; vascular endothelium; NuV129.
XX
OS Homo sapiens.
XX
PN WO9932143-A1.
XX
PD 01-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US27498.
XX
PR 23-DEC-1997; 97US-0996744.
XX
PA (NUVA-) NUVAS LLC.

```
XX
PI   Dickinson CD,  Houston LL;
XX
DR   WPI; 1999-405116/34.
DR   N-PSDB; AAX78893.
XX
PT   New thrombogenic polypeptides used to, e.g. obliterate vasculative
PT   malformations
XX
PS   Claim 30; Page 85-88; 97pp; English.
XX
CC   This invention describes novel thrombogenic polypeptides which comprise a
CC   thrombogenic substructure and a context-dependent entity which recognizes
CC   desired biologically susceptible sites, e.g. tumour vascular endothelium.
CC   A novel context-dependent functional entity comprises a substructure with
CC   thrombogenic potential and one or more context-enhancing substructures
CC   having the ability to recognize desired biologically susceptible sites,
CC   where the entity imparts thrombogenic activity when positioned in the
CC   function-forming-context at the biologically susceptible sites, and the
CC   entity has no thrombogenic activity absent a function-forming-context at
CC   the biologically susceptible sites. The context-dependent functional
CC   entities impart thrombogenic activity only at biologically susceptible
CC   sites. They can be used to obliterate vasculative malformations or to
CC   selectively thrombose the vasculature of solid tumours. This sequence
CC   represents the human tissue factor protein NuV129 which is used in the
CC   method of the invention.
XX
SQ   Sequence   357 AA;

  Query Match             97.1%;  Score 134;  DB 20;  Length 357;
  Best Local Similarity   95.7%;  Pred. No. 2.2e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      96 RNPDGDVGGPWCYTTNPRKLYDY 118


RESULT 23
AAR22502
ID   AAR22502 standard; Protein; 371 AA.
XX
AC   AAR22502;
XX
DT   25-AUG-1992  (first entry)
XX
DE   [GARSYQ]-[Plasminogen 443-541]-[t-PA 262-527] hybrid.
XX
KW   Plasminogen; tissue plasminogen activator; hybrid; fibrinolysis;
KW   blood clotting; acute myocardial infarction.
XX
FH   Key             Location/Qualifiers
FT   Region          1..6
FT                   /note= "amino acids -3 to +3 of t-PA"
FT   Region          7..105
FT                   /note= "amino acids 443-541 of plasminogen"
FT   Region          106..371
FT                   /note= "amino acids 262-527 of t-PA"
FT   Cleavage-site   119..120
FT                   /note= "t-PA cleavage site"
XX
PN   WO9204450-A.
XX
PD   19-MAR-1992.
XX
PF   29-AUG-1991;   91WO-GB01455.
XX
PR   01-SEP-1990;   90GB-0019120.
XX
PA   (BEEC ) BEECHAM GROUP PLC.
XX
PI   Dodd I,  Brown M,  Robinson JH;
XX
DR   WPI; 1992-114357/14.
XX
PT   Hybrid plasminogen activators for treating thrombotic diseases -
PT   comprise Kringle 5 or Kringles 4 and 5 of plasminogen linked to
PT   B-chain of t-PA or u-PA via aminoacid sequence
XX
PS   Claim 9; Page 50; 64pp; English.
XX
CC   This hybrid plasminogen/t-PA sequence is described in the
CC   specification although the sequence itself is not given. The
CC   sequence given here has been compiled using the human plasminogen
CC   sequence from Patent No. WO9013640 and the t-PA sequence of DE3930099.
CC   The variant 4-anisoyl [SYQ]-[plasminogen 443-541]-[t-PA 262-527]
CC   is also specifically claimed. The hybrid fibrinolytic enzymes are
CC   useful in the treatment of thrombotic diseases.
CC   See also AAR22499 and AAR22503-4.
XX
SQ   Sequence   371 AA;

  Query Match             97.1%;  Score 134;  DB 13;  Length 371;
  Best Local Similarity   95.7%;  Pred. No. 2.2e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      77 RNPDGDVGGPWCYTTNPRKLYDY 99


RESULT 24
AAR22504
ID   AAR22504 standard; Protein; 380 AA.
XX
AC   AAR22504;
XX
DT   25-AUG-1992  (first entry)
XX
DE   [GARSYQ]-[Plasminogen 443-541]-[u-PA 137-411] hybrid.
XX
KW   Plasminogen; urokinase-like plasminogen activator; hybrid;
KW   fibrinolysis; blood clotting; acute myocardial infarction.
XX
FH   Key             Location/Qualifiers
FT   Region          1..6
FT                   /note= "amino acids -3 to +3 of t-PA"
FT   Region          7..105
FT                   /note= "amino acids 443-541 of plasminogen"
FT   Region          106..380
FT                   /note= "amino acids 137-411 of u-PA"
FT   Cleavage-site   126..127
FT                   /note= "u-PA cleavage site"
XX
PN   WO9204450-A.
XX
PD   19-MAR-1992.
XX
PF   29-AUG-1991;   91WO-GB01455.
XX
PR   01-SEP-1990;   90GB-0019120.
XX
PA   (BEEC ) BEECHAM GROUP PLC.
XX
PI   Dodd I,  Brown M,  Robinson JH;
XX
DR   WPI; 1992-114357/14.
XX
PT   Hybrid plasminogen activators for treating thrombotic diseases -
PT   comprise Kringle 5 or Kringles 4 and 5 of plasminogen linked to
PT   B-chain of t-PA or u-PA via aminoacid sequence
XX
PS   Claim 11; Page 50; 64pp; English.
XX
CC   This hybrid plasminogen/u-PA sequence is described in the
CC   specification although the sequence itself is not given. The
```

```
CC   sequence given here has been compiled using the human plasminogen
CC   sequence from Patent No. WO9013640 and the human u-PA sequence
CC   from the SWISSPROT database (UROK$HUMAN). The hybrid fibrinolytic
CC   enzymes are useful in the treatment of thrombotic diseases.
CC   See also R22499, R22502 and R22503.
XX
SQ   Sequence   380 AA;

  Query Match             97.1%;  Score 134;  DB 13;  Length 380;
  Best Local Similarity   95.7%;  Pred. No. 2.3e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db     77 RNPDGDVGGPWCYTTNPRKLYDY 99


RESULT 25
AAW51457
ID   AAW51457 standard; protein; 437 AA.
XX
AC   AAW51457;
XX
DT   02-SEP-1998  (first entry)
XX
DE   Human plasminogen fragment with neovascularisation inhibiting activity.
XX
KW   Human plasminogen; neovascularisation; angiostatin; inhibition;
KW   elastase; Sepharose.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  4..81
FT                   /label= Disulphide_bond
FT   Disulfide-bond  25..64
FT                   /label= Disulphide_bond
FT   Disulfide-bond  53..76
FT                   /label= Disulphide_bond
FT   Disulfide-bond  108..187
FT                   /label= Disulphide_bond
FT   Disulfide-bond  129..170
FT                   /label= Disulphide_bond
FT   Disulfide-bond  158..182
FT                   /label= Disulphide_bond
FT   Disulfide-bond  194..313
FT                   /label= Disulphide_bond
FT   Disulfide-bond  204..212
FT                   /label= Disulphide_bond
FT   Disulfide-bond  234..250
FT                   /label= Disulphide_bond
FT   Disulfide-bond  326..393
FT                   /label= Disulphide_bond
FT   Disulfide-bond  356..372
FT                   /label= Disulphide_bond
FT   Disulfide-bond  383..411
FT                   /label= Disulphide_bond
XX
PN   JP10158300-A.
XX
PD   16-JUN-1998.
XX
PF   28-NOV-1996;   96JP-0317250.
XX
PR   28-NOV-1996;   96JP-0317250.
XX
PA   (SUZM ) SUZUKI KK.
XX
DR   WPI; 1998-393476/34.
XX
PT   Human plasminogen derived polypeptide - has neovascularisation
PT   inhibiting activity
XX
PS   Claim 1; Page 2; 16pp; Japanese.
XX
CC   The invention relates to a neovascularisation inhibitor which comprises
CC   amino acids 355-791 of human plasminogen. Also claimed are a method for
CC   the preparation of angiostatin, and angiostatin prepared by this method.
CC   The human plasminogen protein fragment is prepared by: (a) applying human
CC   plasminogen to a lysine Sepharose column to separate it into plasminogen
CC   form 1 and form 2; (b) separating plasminogen form 1 and form 2 and
CC   digesting them with elastase; (c) fractionating the elastase-decomposed
CC   product of form 1 plasminogen and form 2 plasminogen in a lysine
CC   Sepharose column; (d) collecting the fractions bound to the lysine
CC   Sepharose column; (e) further fractionating the form 2 plasminogen using
CC   an Aminohexal Sepharose column; and (f) collecting the fraction bound to
CC   the Aminohexal Sepharose column. This human plasminogen fragment can be
CC   used to inhibit growth of vascular endothelial cells. The present
CC   sequence represents amino acids 355-791 of human plasminogen.
XX
SQ   Sequence   437 AA;

  Query Match             97.1%;  Score 134;  DB 19;  Length 437;
  Best Local Similarity   95.7%;  Pred. No. 2.6e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db    159 RNPDGDVGGPWCYTTNPRKLYDY 181


RESULT 26
AAR22499
ID   AAR22499 standard; protein; 467 AA.
XX
AC   AAR22499;
XX
DT   25-AUG-1992  (first entry)
XX
DE   [GARSYQ]-[Plasminogen 347-541]-[t-PA 262-527] hybrid.
XX
KW   Plasminogen; tissue plasminogen activator; hybrid; fibrinolysis;
KW   blood clotting; acute myocardial infarction.
XX
FH   Key             Location/Qualifiers
FT   Region          1..6
FT                   /note= "amino acids -3 to +3 of t-PA"
FT   Region          7..201
FT                   /note= "amino acids 347-541 of plasminogen"
FT   Region          202..467
FT                   /note= "amino acids 262-527 of t-PA"
FT   Cleavage-site   215..216
FT                   /note= "t-PA cleavage site"
XX
PN   WO9204450-A.
XX
PD   19-MAR-1992.
XX
PF   29-AUG-1991;   91WO-GB01455.
XX
PR   01-SEP-1990;   90GB-0019120.
XX
PA   (BEEC ) BEECHAM GROUP PLC.
XX
PI   Dodd I,  Brown M,  Robinson JH;
XX
DR   WPI; 1992-114357/14.
XX
PT   Hybrid plasminogen activators for treating thrombotic diseases -
PT   comprise Kringle 5 or Kringles 4 and 5 of plasminogen linked to
PT   B-chain of t-PA or u-PA via aminoacid sequence
XX
PS   Claim 8; Page 50; 64pp; English.
XX
CC   This hybrid plasminogen/t-PA sequence is described in the
CC   specification although the sequence itself is not given. The
```

```
CC   sequence given here has been compiled using the human plasminogen
CC   sequence from Patent No. WO9013640 and the t-PA sequence of DE3930099.
CC   A N,N-Dimethyl-4-aminobenzoyl two-chain [SYQ]-[plasminogen 347-541]-
CC   [t-PA 262-527] and the 4-anisoyl version of this variant are also
CC   specifically claimed. The hybrid fibrinolytic enzymes are useful in
CC   the treatment of thrombotic diseases. See also AAR22502-4.
XX
SQ   Sequence   467 AA;

  Query Match             97.1%;  Score 134;  DB 13;  Length 467;
  Best Local Similarity   95.7%;  Pred. No. 2.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     173 RNPDGDVGGPWCYTTNPRKLYDY 195


RESULT 27
AAR22503
ID   AAR22503 standard; Protein; 476 AA.
XX
AC   AAR22503;
XX
DT   25-AUG-1992  (first entry)
XX
DE   [GARSYQ]-[Plasminogen 347-541]-[u-PA 137-411] hybrid.
XX
KW   Plasminogen; urokinase-like plasminogen activator; hybrid;
KW   fibrinolysis; blood clotting; acute myocardial infarction.
XX
FH   Key             Location/Qualifiers
FT   Region          1..6
FT                   /note= "amino acids -3 to +3 of t-PA"
FT   Region          7..201
FT                   /note= "amino acids 347-541 of plasminogen"
FT   Region          202..476
FT                   /note= "amino acids 137-411 of u-PA"
FT   Cleavage-site   222..223
FT                   /note= "u-PA cleavage site"
XX
PN   WO9204450-A.
XX
PD   19-MAR-1992.
XX
PF   29-AUG-1991;   91WO-GB01455.
XX
PR   01-SEP-1990;   90GB-0019120.
XX
PA   (BEEC ) BEECHAM GROUP PLC.
XX
PI   Dodd I,  Brown M,  Robinson JH;
XX
DR   WPI; 1992-114357/14.
XX
PT   Hybrid plasminogen activators for treating thrombotic diseases -
PT   comprise Kringle 5 or Kringles 4 and 5 of plasminogen linked to
PT   B-chain of t-PA or u-PA via aminoacid sequence
XX
PS   Claim 10; Page 50; 64pp; English.
XX
CC   This hybrid plasminogen/u-PA sequence is described in the
CC   specification although the sequence itself is not given. The
CC   sequence given here has been compiled using the human plasminogen
CC   sequence from Patent No. WO9013640 and the human u-PA sequence
CC   from the SWISSPROT database (UROK$HUMAN). The hybrid fibrinolytic
CC   enzymes are useful in the treatment of thrombotic diseases.
CC   See also R22499, R22502 and R22504.
XX
SQ   Sequence   476 AA;

  Query Match             97.1%;  Score 134;  DB 13;  Length 476;
  Best Local Similarity   95.7%;  Pred. No. 2.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     173 RNPDGDVGGPWCYTTNPRKLYDY 195


RESULT 28
AAY02100
ID   AAY02100 standard; Protein; 566 AA.
XX
AC   AAY02100;
XX
DT   16-JUL-1999  (first entry)
XX
DE   A multifunctional protein of the invention.
XX
KW   Angiostatin; endostatin; interferon; thrombospondin;
KW   interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW   anti-tumor;. multifunctional protein; angiogenic-mediated disease;
KW   cancer; diabetic retinopathy; macular degeneration; arthritis;
KW   tumor cell production.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916889-A1.
XX
PD   08-APR-1999.
XX
PF   30-SEP-1998;   98WO-US20464.
XX
PR   01-OCT-1997;   97US-0060609.
XX
PA   (SEAR ) SEARLE & CO G D.
XX
PI   Bolanowski MA,  Caparon MH,  Casperson GF,  Gregory SA;
PI   Klein BK,  McKearn JP;
XX
DR   WPI; 1999-255098/21.
XX
PT   New multifunctional proteins useful for treating angiogenic-mediated
PT   diseases
XX
PS   Claim 5; Page 95-96; 121pp; English.
XX
CC   The specification describes multifunctional proteins which comprise
CC   combinations of angiostatin, endostatin, interferon, thrombospondin,
CC   interferon-inducible protein and platelet factor 4, and have
CC   anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC   may exhibit useful properties such as having similar or greater
CC   biological activity when compared to a single factor or by having
CC   improved half-life or decreased adverse side effects, or a combination
CC   of these properties. The proteins can be used for treating an
CC   angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC   degeneration, or arthritis. They can also be used for inhibiting the
CC   production of tumor cells (characteristic of lung, breast, ovarian,
CC   prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC   hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC   growth. The present sequence represents a multifunctional protein of the
CC   invention.
XX
SQ   Sequence   566 AA;

  Query Match             97.1%;  Score 134;  DB 20;  Length 566;
  Best Local Similarity   95.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554
```

RESULT 29
AAR60519
ID AAR60519 standard; Protein; 790 AA.
XX
AC AAR60519;
XX
DT 22-MAR-1995 (first entry)
XX
DE Human 'Glu' plasminogen.
XX
KW Serine protease; Factor-Xa; recognition site; plasminogen; kringle;
KW fusion protein cleavage; protein folding; primer;
KW polymerase chain reaction; amplification.
XX
OS Homo sapiens.
XX
PN WO9418227-A.
XX
PD 18-AUG-1994.
XX
PF 04-FEB-1994; 94WO-DK00054.
XX
PR 04-FEB-1993; 93DK-0000130.
PR 05-FEB-1993; 93DK-0000139.
PR 03-DEC-1993; 93WO-GB02492.
XX
PA (DENZ-) DENZYME APS.
XX
PI Etzerodt M, Holtet TL, Thogersen HC;
XX
DR WPI; 1994-279681/34.
XX
PT Refolding of polypeptide molecules - using a cyclic process
PT involving denaturing and renaturing conditions to produce a
PT correctly folded prod
XX
PS Disclosure; Page 148-50; 202pp; English.
XX
CC cDNA encoding kringle domains 1 and 4 of human plasminogen (full
CC sequence given in AAR60519) was PCR amplified using primers given in
CC AAQ71268-71. Amplified cDNA was linked to a sequence encoding the
CC Factor-Xa cleavage site (given in AAR60503), subcloned in vector
CC pLcIIMLCH6 so that it was linked to a hexahistidine-encoding
CC sequence and expressed in E. coli QY13. The fusion protein was
CC purified on an Ni2+-activated NTA-agarose column. A cyclic
CC procedure was used to obtain correctly folded recombinant protein.
XX
SQ Sequence 790 AA;

Query Match 97.1%; Score 134; DB 15; Length 790;
Best Local Similarity 95.7%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 512 RNPDGDVGGPWCYTTNPRKLYDY 534

RESULT 30
AAB36562
ID AAB36562 standard; protein; 790 AA.
XX
AC AAB36562;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mammalian kringle 5 protein SEQ ID NO:1.
XX
KW Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
KW vasotropic; cytostatic; antirheumatic; antipsoriatic; antidiabetic;
KW antiarteriosclerotic; osteopathic; angiogenesis inhibitor; angiogenesis;
KW inflammatory disorder; inflammation; chronic articular rheumatism;
KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
KW capillary proliferation; atherosclerotic plaque; osteoporosis;
KW cancer; solid tumour; angiofibroma; retrolental fibroplasia;
KW haemangioma; Kasposi's sarcoma; neovascularisation; tumour growth.
XX
OS Mammalia.
XX
PN WO200070665-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-IB00763.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Rasamoelisolo M, Thibaudeau K, Huang X, Beliveau R;
XX
DR WPI; 2001-090970/10.
XX
PT New modified anti-angiogenic kringle 5 peptides capable of forming
PT conjugates with blood proteins, useful for treating angiogenesis,
PT inappropriate invasion of vessels or cancers in humans or mammals -
XX
PS Disclosure; Page 74-77; 82pp; English.
XX
CC The present invention describes a modified anti-angiogenic peptide (I)
CC comprising a reactive group that reacts with amino groups, hydroxyl
CC groups or thiol groups on blood components to form stable covalent
CC bonds. The reactive group is selected from succinimidyl or maleimido
CC groups. (I) can have anti-inflammatory, vasotropic, cytostatic,
CC antirheumatic, antipsoriatic, antidiabetic, antiarteriosclerotic and
CC osteopathic activities, and is an angiogenesis inhibitor. (I) are useful
CC for treating angiogenesis in a human, where the derivative is reacted
CC with blood proteins. (I) are also useful for manufacturing a medicament
CC extending the in vivo half-life of a kringle 5 peptide in a patient to
CC provide an anti-angiogenic effect. In particular, a modified kringle 5
CC peptide can be used for treating inflammatory disorders (e.g. immune and
CC non-immune inflammation, chronic articular rheumatism or psoriasis),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (e.g. diabetic retinopathy, neovascular glaucoma, restenosis,
CC capillary proliferation in atherosclerotic plaques or osteoporosis), or
CC cancer associated disorders (e.g. solid tumours, solid tumour
CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas,
CC Kasposi's sarcoma or other cancers requiring neovascularisation to
CC support tumour growth). The peptides are useful for treating these
CC diseases in mammalian or human patients. AAB36562 represents a mammalian
CC kringle 5 protein, and AAB36563 to AAB36577 represent specifically
CC claimed kringle 5 peptides from the present invention.
XX
SQ Sequence 790 AA;

Query Match 97.1%; Score 134; DB 22; Length 790;
Best Local Similarity 95.7%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 512 RNPDGDVGGPWCYTTNPRKLYDY 534

RESULT 31
AAB01887
ID AAB01887 standard; Protein; 791 AA.
XX
AC AAB01887;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen, SEQ ID NO:1.
XX
KW Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;

```
KW   antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW   antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  462..541
FT   Disulfide-bond  483..524
FT   Disulfide-bond  512..536
FT   Misc-difference 682
FT                   /note= "Encoded by GTC"
XX
PN   US6057122-A.
XX
PD   02-MAY-2000.
XX
PF   05-MAY-1997;   97US-0851350.
XX
PR   03-MAY-1996;   96US-0643219.
PR   03-APR-1997;   97US-0832087.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davidson DJ;
XX
DR   WPI; 2000-349573/30.
DR   N-PSDB; AAA52284.
XX
PT   Preparation of Kringle five peptide fragment for treating various
PT   disorders such as angiogenic, ocular, skin diseases and cancer,
PT   involves mixing mammalian plasminogen and elastase followed by
PT   incubation and isolation -
XX
PS   Example 17; Fig 1; 48pp; English.
XX
CC   The invention relates to a method of preparing plasminogen kringle 5
CC   peptide fragments. The method comprises mixing mammalian plasminogen and
CC   elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC   the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC   endothelial cell proliferation and migration. The peptides are useful for
CC   treating angiogenic diseases, primary and metastatic solid tumours and
CC   carcinomas of various organs such as breast, genital tract, endocrine
CC   glands, skin, tumours of the brain and eyes and solid tumours arising
CC   from haematopoietic malignancies such as leukaemias and lymphomas. They
CC   are also used for the prophylaxis of various autoimmune diseases (e.g.,
CC   rheumatoid arthritis), ocular diseases, skin diseases (e.g., psoriasis),
CC   blood vessel diseases (e.g. haemangiomas, Osler-Webber Syndrome),
CC   diseases caused by excessive or abnormal stimulation of endothelial cells
CC   (e.g., Crohn's disease, atherosclerosis), diseases which have
CC   angiogenesis as a pathologic consequence (e.g., cat scratch disease and
CC   ulcers). The peptides are also useful as a birth control agent which
CC   inhibits ovulation and establishment of the placenta. The present
CC   sequence represents human plasminogen.
XX
SQ   Sequence   791 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 791;
  Best Local Similarity   95.7%;  Pred. No. 4.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     513 RNPDGDVGGPWCYTTNPRKLYDY 535


RESULT 32
AAY99589
ID   AAY99589 standard; protein; 791 AA.
XX
AC   AAY99589;
XX
DT   13-SEP-2000  (first entry)
XX
DE   Human plasminogen.
XX
KW   Human; serine protease zymogen; cardiant; thrombolytic;
KW   plasminogen activator; heart attack; stroke; blood clotting disorder.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 644..645
FT                   /note= "specifically claimed in claim 6 of the
FT                   specification"
FT   Misc-difference 542..791
FT                   /note= "specifically claimed in claims 7 and 8 of the
FT                   specification, preferably amino acids 562..791
FT                   (claim 7)"
XX
PN   WO200032759-A1.
XX
PD   08-JUN-2000.
XX
PF   06-MAY-1999;   99WO-US09991.
XX
PR   02-DEC-1998;   98US-0110588.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   Lin X,  Zhang XC,  Tang JJN;
XX
DR   WPI; 2000-422975/36.
XX
PT   Polypeptide with plasminogen activator activity useful as thrombolytic
PT   agent for treating blood clot disorders e.g. heart attack, comprises 10
PT   amino acid peptide fragment for recognition or activation of
PT   plasminogen  -
XX
PS   Claim 6,7,8; Page 20-23; 41pp; English.
XX
CC   The present sequence is plasminogen, the principal serine
CC   protease zymogen in the extracellular fluids of vertebrates. Its
CC   active form, plasmin, is implicated in pericellular proteolysis
CC   associated with a wide range of physiological and pathological
CC   processes. Plasminogen expression is regulated by plasminogen activators
CC   which hydrolyse a peptide bond in plasminogen to convert it to plasmin
CC   or form tight binding complexes with plasminogen to spontaneously
CC   convert it to plasmin. Review of sequence homologies of several
CC   plasminogen activators and chymotrypsin has identified a six amino acid
CC   peptide involved in plasminogen activation. This peptide is particularly
CC   useful when inserted between amino acid residues 644 and 645 of full
CC   length human plasminogen. Novel plasminogen activators have been made
CC   based upon the plasminogen activation/recognition site of plasminogen
CC   binding proteins. The polypeptides are useful in preparing thrombolytic
CC   agents for treating blood clotting disorders such as heart attack.
XX
SQ   Sequence   791 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 791;
  Best Local Similarity   95.7%;  Pred. No. 4.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     513 RNPDGDVGGPWCYTTNPRKLYDY 535


RESULT 33
AAY50867
ID   AAY50867 standard; protein; 791 AA.
XX
AC   AAY50867;
XX
DT   24-FEB-2000  (first entry)
XX
DE   Human plasminogen protein fragment.
```

```
XX
KW   Plasminogen; human; thrombolytic agent; streptokinase; antigenic;
KW   blood clot; heart attack; treatment.
XX
OS   Homo sapiens.
XX
PN   WO9957251-A2.
XX
PD   11-NOV-1999.
XX
PF   06-MAY-1999;   99WO-US10086.
XX
PR   06-MAY-1998;   98US-0084392.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   Zhang XC,  Lin X,  Tang JJN;
XX
DR   WPI; 2000-052966/04.
XX
PT   New thrombolytic agents derived from modified humanized.streptokinase,
PT   useful for treating blood clot disorders -
XX
PS   Example 5; Page 40-43; 55pp; English.
XX
CC   This invention describes a novel thrombolytic agent comprising
CC   streptokinase where at least one nonessential portion has been modified.
CC   The invention also describes a method of forming a thrombolytic agent
CC   comprising determining a nonessential portion of streptokinase and
CC   modifying the nonessential portion to render the resulting protein less
CC   antigenic. The modified streptokinase is used to treat blood clot
CC   disorders, such as heart attacks. The modified streptokinase has less
CC   antigenicity than streptokinase but is still able to complex plasminogen
CC   and lead to plasminogen activation. Modified streptokinase with the
CC   nonessential portions removed or truncated simplify the molecule. Such
CC   smaller proteins are cheaper and easier to produce. This sequence
CC   represents a fragment of the human plasminogen protein which is used in
CC   the description of the method of the invention.
XX
SQ   Sequence   791 AA;

  Query Match             97.1%;  Score 134;  DB 21;  Length 791;
  Best Local Similarity   95.7%;  Pred. No. 4.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     513 RNPDGDVGGPWCYTTNPRKLYDY 535


RESULT 34
AAG67223
ID   AAG67223 standard; Protein; 791 AA.
XX
AC   AAG67223;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Amino acid sequence of human plasminogen.
XX
KW   Angiostatin; plasminogen; sulfydryl donor; angiogenesis; tumour;
KW   angiogenic disease; neoplastic disease; connective tissue disorder;
KW   rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;
KW   diabetic retinopathy; corneal graft rejection; cardiovascular disease;
KW   cerebral vascular disease; diabetes; immune disorder;
KW   chronic inflammation; autoimmunity.
XX
OS   Homo sapiens.
XX
PN   WO200158921-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-US04021.
XX
PR   08-FEB-2000; 2000US-0500397.
XX
PA   (NOUN ) UNIV NORTHWESTERN.
XX
PI   Soff G,  Gately ST,  Twardowski P;
XX
DR   WPI; 2001-550019/61.
XX
PT   Producing angiostatin for treating angiogenic diseases involves
PT   contacting plasminogen with plasminogen activator and sulfhydryl donor
PT   simultaneously, or producing plasmin which is contacted with sulfhydryl
PT   donor -
XX
PS   Disclosure; Page 77-80; 101pp; English.
XX
CC   The specification describes a method for generating angiostatin in
CC   vitro. The method comprises contacting plasminogen with a sulfydryl
CC   donor, or culturing cells capable of producing plasminogen activator
CC   in conditioned culture medium (CCM) and contacting the CCM with
CC   plasminogen. Angiostatin produced by method of the invention is useful
CC   for treating animals with angiogenesis diseases. It is useful for
CC   treating an angiogenic disease such as neoplastic diseases (e.g. tumours
CC   and tumour metastasis), benign tumours (e.g. hemangiomas, acoustic
CC   neuromas, etc), connective tissue disorders (e.g. rheumatoid arthritis
CC   and atherosclerosis), ocular angiogenic diseases (e.g. diabetic
CC   retinopathy, corneal graft rejection, etc), cardiovascular diseases,
CC   cerebral vascular diseases, diabetes-associated diseases and immune
CC   disorders (e.g. chronic inflammation and autoimmunity). The present
CC   sequence represents a plasminogen.
XX
SQ   Sequence   791 AA;

  Query Match             97.1%;  Score 134;  DB 22;  Length 791;
  Best Local Similarity   95.7%;  Pred. No. 4.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     513 RNPDGDVGGPWCYTTNPRKLYDY 535


RESULT 35
AAR20013
ID   AAR20013 standard; Protein; 807 AA.
XX
AC   AAR20013;
XX
DT   31-MAR-1992  (first entry)
XX
DE   PA mutant Plg 1-541 [Arg298-299->Gln298-299] t-PA 262-527.
XX
KW   Fibrinolytic; plasminogen activator; kringle; t-PA; thrombosis;
KW   myocardial infarction; cleavage site.
XX
FH   Key             Location/Qualifiers
FT   Region          1..541
FT                   /label= plasminogen
FT                   /note= "residues 1-541, including kringle
FT                         domains 1-5"
FT   Region          542..807
FT                   /label= t-PA
FT                   /note= "[Arg298->Gln, Arg299->Gln]t-PA 262-527"
XX
PN   WO9118989-A.
XX
PD   12-DEC-1991.
XX
PF   21-MAY-1991;   91WO-GB00801.
XX
PR   26-MAY-1990;   90GB-0011861.
XX
```

```
PA   (BEEC ) BEECHAM GROUP PLC.
XX
PI   Dodd I;
XX
DR   WPI; 1992-007472/01.
XX
PT   New hybrid plasminogen activator mutants - useful in treating
PT   thrombotic diseases esp. myocardial infarction
XX
PS   Claim 6; Page 30; 31pp; English.
XX
CC   The fibrinolytic active plasminogen activator (PA) comprises the five
CC   kringle domains of plasminogen linked via an amino acid sequence
CC   comprising the t-PA cleavage site between residues 275 and 276
CC   and Cys264 of t-PA, to a t-PA B-chain modified to have a residue
CC   other than Arg at position 299.  One and two chain variants, Lys78
CC   and Glu1 variants and mixt. of these are included.
CC   The PA is used on compsns. to treat thrombotic diseases, in partic.
CC   myocardial infarction.  Dosage is 0.1-10 (pref. 0.1-2.0) mg/kg by
CC   injection for up to 5 doses, or by infusion.
XX
SQ   Sequence   807 AA;

  Query Match             97.1%;  Score 134;  DB 13;  Length 807;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     513 RNPDGDVGGPWCYTTNPRKLYDY 535


RESULT 36
AAR08065
ID   AAR08065 standard; protein; 810 AA.
XX
AC   AAR08065;
XX
DT   26-FEB-1991  (first entry)
XX
DE   Human plasminogen from plasmid p119PN127-6.
XX
KW   Human plasminogen; thrombolytic therapy; p119PN127.6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= LEADER_SEQUENCE
FT   Protein         20..810
FT                   /label= MATURE_PROTEIN
FT   Domain          103..182
FT                   /label= KRINGLE_1
FT   Domain          186..263
FT                   /label= KRINGLE_2
FT   Domain          275..353
FT                   /label= KRINGLE_3
FT   Domain          377..455
FT                   /label= KRINGLE_4
FT   Domain          479..561
FT                   /label= KRINGLE 5
FT   Protein         561..810
FT                   /label= PMGN PROTEIN PROTEASE
XX
PN   WO9013640-A.
XX
PD   15-NOV-1990.
XX
PF   26-APR-1990;   90WO-US02296.
XX
PR   01-MAY-1989;   89US-0345801.
XX
PA   (UYNO-) UNIV NOTRE DAME DU.
XX
PI   Castellino FJ,  whitefleet JL,  Rosen ED,  Mclinden JH;
XX
DR   WPI; 1990-361475/48.
DR   N-PSDB; AAQ06648.
XX
PT   Expression of human plasminogen in eukaryotic cell vector - used
PT   in thrombolytic therapy
XX
PS   Disclosure; Fig 2; 110pp; English.
XX
CC   Plasminogen can be produced by culturing a eukaryotic cell lacking
CC   a site-specific plasminogen activator and comprising the DNA
CC   encoding this amino acid sequence.
CC   Pharmaceutical compsns. containing plasminogen can be used for
CC   thrombolytic therapy.
XX
SQ   Sequence   810 AA;

  Query Match             97.1%;  Score 134;  DB 11;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 37
AAR13219
ID   AAR13219 standard; Protein; 810 AA.
XX
AC   AAR13219;
XX
DT   05-SEP-1991  (first entry)
XX
DE   R561G human plasminogen variant.
XX
KW   Plasminogen; proteolytic cleavage; variant; thrombosis; HPg.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= sig_peptide
FT   Protein         20..810
FT                   /label= mat_protein
FT   Domain          103..182
FT                   /label= kringle_1
FT   Domain          185..263
FT                   /label= kringle_2
FT   Domain          275..353
FT                   /label= kringle_3
FT   Domain          377..455
FT                   /label= kringle_4
FT   Domain          481..561
FT                   /label= kringle_5
FT   Region          561..810
FT                   /label= protein_protease
XX
PN   WO9108297-A.
XX
PD   13-JUN-1991.
XX
PF   31-OCT-1990;   90WO-US06345.
XX
PR   01-DEC-1989;   89US-0444584.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Castellino FJ,  Higgins D L;
XX
DR   WPI; 1991-193201/26.
```

```
DR   N-PSDB; AAQ11998.
XX
PT   New human plasminogen variant with replaced ARG-561 - is
PT   complexed with fibrinolytic enzyme for use in thrombolytic
PT   therapy
XX
PS   Disclosure; Fig 1(A-P); 82pp; English.
XX
CC   The amino acid at position 475 is Val for the vector pUC119PN127.6,
CC   whereas Fig 1 of the specification reflecting the sequence
CC   encoding native human plasminogen, contains Ala.  The amino acid at
CC   561 of the wild-type plasminogen (Arg), is substituted by any amino
CC   acid except Lys.  Here the amino acids is substituted by Gly.
CC   Position 561 is the critical cleavage site in the conversion of
CC   plasminogen to plasmin.  The resulting product is proteolytic
CC   resistant plasminogen which may be used to treat thrombosis in
CC   humans.  When complexed with streptokinase, it does not degrade into
CC   plasmin and it caused rapid plasminogen activator activity.
XX
SQ   Sequence   810 AA;

  Query Match             97.1%;  Score 134;  DB 12;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 38
AAR12406
ID   AAR12406 standard; Protein; 810 AA.
XX
AC   AAR12406;
XX
DT   05-SEP-1991  (first entry)
XX
DE   R561E human plasminogen variant.
XX
KW   Plasminogen; proteolytic cleavage; variant; thrombosis; HPg.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= sig_peptide
FT   Protein         20..810
FT                   /label= mat_protein
FT   Domain          103..182
FT                   /label= kringle_1
FT   Domain          185..263
FT                   /label= kringle_2
FT   Domain          275..353
FT                   /label= kringle_3
FT   Domain          377..455
FT                   /label= kringle_4
FT   Domain          481..561
FT                   /label= kringle_5
FT   Region          561..810
FT                   /label= protein_protease
XX
PN   WO9108297-A.
XX
PD   13-JUN-1991.
XX
PF   31-OCT-1990;   90WO-US06345.
XX
PR   01-DEC-1989;   89US-0444584.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Castellino FJ,  Higgins D L;
XX
DR   WPI; 1991-193201/26.
DR   N-PSDB; AAQ11998.
XX
PT   New human plasminogen variant with replaced ARG-561 - is
PT   complexed with fibrinolytic enzyme for use in thrombolytic
PT   therapy
XX
PS   Disclosure; Fig 1(A-P); 82pp; English.
XX
CC   The amino acid at position 475 is Val for the vector pUC119PN127.6,
CC   whereas Fig 1 of the specification reflecting the sequence
CC   encoding native human plasminogen, contains Ala.  The amino acid at
CC   561 of the wild-type plasminogen (Arg), is substituted by any amino
CC   acid except Lys.  Here the amino acids is substituted by Glu.
CC   Position 561 is the critical cleavage site in the conversion of
CC   plasminogen to plasmin.  The resulting product is proteolytic
CC   resistant plasminogen which may be used to treat thrombosis in
CC   humans.  When complexed with streptokinase, it does not degrade into
CC   plasmin and it caused rapid plasminogen activator activity.
XX
SQ   Sequence   810 AA;

  Query Match             97.1%;  Score 134;  DB 12;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 39
AAR13220
ID   AAR13220 standard; Protein; 810 AA.
XX
AC   AAR13220;
XX
DT   05-SEP-1991  (first entry)
XX
DE   R561S human plasminogen variant.
XX
KW   Plasminogen; proteolytic cleavage; variant; thrombosis; HPg.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= sig_peptide
FT   Protein         20..810
FT                   /label= mat_protein
FT   Domain          103..182
FT                   /label= kringle_1
FT   Domain          185..263
FT                   /label= kringle_2
FT   Domain          275..353
FT                   /label= kringle_3
FT   Domain          377..455
FT                   /label= kringle_4
FT   Domain          481..561
FT                   /label= kringle_5
FT   Region          561..810
FT                   /label= protein_protease
XX
PN   WO9108297-A.
XX
PD   13-JUN-1991.
XX
PF   31-OCT-1990;   90WO-US06345.
XX
PR   01-DEC-1989;   89US-0444584.
XX
PA   (GETH ) GENENTECH INC.
```

```
XX
PI   Castellino FJ,  Higgins D L;
XX
DR   WPI; 1991-193201/26.
DR   N-PSDB; AAQ11998.
XX
PT   New human plasminogen variant with replaced ARG-561 - is
PT   complexed with fibrinolytic enzyme for use in thrombolytic
PT   therapy
XX
PS   Disclosure; Fig 1(A-P); 82pp; English.
XX
CC   The amino acid at position 475 is Val for the vector pUC119PN127.6,
CC   whereas Fig 1 of the specification reflecting the sequence
CC   encoding native human plasminogen, contains Ala.  The amino acid at
CC   561 of the wild-type plasminogen (Arg), is substituted by any amino
CC   acid except Lys.  Here the amino acids is substituted by Ser.
CC   Position 561 is the critical cleavage site in the conversion of
CC   plasminogen to plasmin.  The resulting product is proteolytic
CC   resistant plasminogen which may be used to treat thrombosis in
CC   humans.  When complexed with streptokinase, it does not degrade into
CC   plasmin and it caused rapid plasminogen activator activity.
XX
SQ   Sequence   810 AA;

  Query Match             97.1%;  Score 134;  DB 12;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| |||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 40
AAR13221
ID   AAR13221 standard; Protein; 810 AA.
XX
AC   AAR13221;
XX
DT   05-SEP-1991  (first entry)
XX
DE   Human plasminogen variant.
XX
KW   Plasminogen; proteolytic cleavage; variant; thrombosis; HPg.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= sig_peptide
FT   Protein         20..810
FT                   /label= mat_protein
FT   Domain          103..182
FT                   /label= kringle_1
FT   Domain          185..263
FT                   /label= kringle_2
FT   Domain          275..353
FT                   /label= kringle_3
FT   Domain          377..455
FT                   /label= kringle_4
FT   Domain          481..561
FT                   /label= kringle_5
FT   Region          561..810
FT                   /label= protein_protease
XX
PN   WO9108297-A.
XX
PD   13-JUN-1991.
XX
PF   31-OCT-1990;   90WO-US06345.
XX
PR   01-DEC-1989;   89US-0444584.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Castellino FJ,  Higgins D L;
XX
DR   WPI; 1991-193201/26.
DR   N-PSDB; AAQ11998.
XX
PT   New human plasminogen variant with replaced ARG-561 - is
PT   complexed with fibrinolytic enzyme for use in thrombolytic
PT   therapy
XX
PS   Disclosure; Fig 1(A-P); 82pp; English.
XX
CC   The amino acid at position 475 is Val for the vector pUC119PN127.6,
CC   whereas Fig 1 of the specification reflecting the sequence
CC   encoding native human plasminogen, contains Ala.  The amino acid at
CC   561 of the wild-type plasminogen (Arg), is substituted by any amino
CC   acid except Lys.
CC   Position 561 is the critical cleavage site in the conversion of
CC   plasminogen to plasmin.  The resulting product is proteolytic
CC   resistant plasminogen which may be used to treat thrombosis in
CC   humans.  When complexed with streptokinase, it does not degrade into
CC   plasmin and it caused rapid plasminogen activator activity.
XX
SQ   Sequence   810 AA;

  Query Match             97.1%;  Score 134;  DB 12;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| |||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 41
AAR12938
ID   AAR12938 standard; Protein; 810 AA.
XX
AC   AAR12938;
XX
DT   23-SEP-1991  (first entry)
XX
DE   Plasminogen mutein T1 with thrombin cleavage site.
XX
KW   protease; fibrinolysis; blood clotting; thrombosis.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   578..580
FT                   /label= Thrombin cleavage site
XX
PN   WO9109118-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1990;   90WO-GO01912.
XX
PR   07-DEC-1989;   89GB-0027722.
PR   07-DEC-1990;   90WO-GB01911.
XX
PA   (BRBI-) BRIT BIO-TECHN LTD.
XX
PI   Dawson KM,  Edwards RM,  Forman JM;
XX
DR   WPI; 1991-208145/28.
DR   N-PSDB; AAQ12547.
XX
PT   Activatable fibrinolytic and antithrombic proteins - activated by
PT   e.g. factor Xa, thrombin or activated protein C
XX
PS   Claim 7; Fig 2 and Fig 5; 73pp; English.
XX
```

```
CC   This protein is a plasminogen mutant cleavable by thrombin.
CC   Activation is localised to the thrombus because cleavage to plasmin
CC   is by an enzyme of the blood clotting pathway. Compositions
CC   comprising the mutant plasminogen are used for treatment or
CC   prevention of thrombosis, etc.
CC   See AAQ12542-Q12558.
XX
SQ   Sequence   810 AA;

  Query Match            97.1%;  Score 134;  DB 12;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| |||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 42
AAR34428
ID   AAR34428 standard; Protein; 810 AA.
XX
AC   AAR34428;
XX
DT   17-AUG-1993  (first entry)
XX
DE   Sequence encoded by a plasminogen cDNA.
XX
KW   Zymogen; fibrinolytic activity; cleavage.
XX
OS   Synthetic.
XX
PN   US5200340-A.
XX
PD   06-APR-1993.
XX
PF   22-MAY-1987;   87US-0053412.
XX
PR   22-MAY-1987;   87US-0053412.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Foster DC,  Mulvihill ER,  Ohara PJ,  Pingel K,  Yoshitake S;
XX
DR   WPI; 1993-133739/16.
DR   N-PSDB; AAQ40319.
XX
PT   Human tissue plasminogen activator single chain form fibrinolytic
PT   agent - comprises thrombin cleavable zymogen stimulating amido
PT   lytic activity, for lysing clots in heart attack and stroke
PT   victims and suppressing fibrin matrix
XX
PS   Example; Fig 10A, 10B, 10C; 22pp; English.
XX
CC   A lambda phage clone comprising a cDNA sequence encoding
CC   plasminogen was obtained from Dr. Mark Martsen at the University of
CC   Washington. The cDNA was isolated from a human liver library by
CC   probing with the partial sequence of Malinowski et al. The sequence
CC   of the complete cDNA and the encoded amino acid sequence are shown
CC   in AAQ40319 and AAR34428.
XX
SQ   Sequence   810 AA;

  Query Match            97.1%;  Score 134;  DB 14;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| |||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 43
AAW31169
ID   AAW31169 standard; Protein; 810 AA.
XX
AC   AAW31169;
XX
DT   26-FEB-1998  (first entry)
XX
DE   Plasminogen protein for production of recombinant plasminogen.
XX
KW   Plasminogen; alpha-1-antitrypsin; AAT; argserpins; recombinant;
KW   stabilising protein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 60
FT                   /note= "encoded by ACC"
FT   Misc-difference 191
FT                   /note= "encoded by GAA"
FT   Misc-difference 223
FT                   /note= "encoded by AAA"
FT   Misc-difference 280
FT                   /note= "encoded by GCT"
FT   Misc-difference 291
FT                   /note= "encoded by CTG"
FT   Misc-difference 395
FT                   /note= "encoded by GSA"
FT   Misc-difference 503
FT                   /note= "encoded by CAG"
FT   Misc-difference 607
FT                   /note= "encoded by TCT"
FT   Misc-difference 615
FT                   /note= "encoded by CAG"
FT   Misc-difference 658
FT                   /note= "encoded by TTG"
FT   Misc-difference 709
FT                   /note= "encoded by GCT"
XX
PN   US5648254-A.
XX
PD   15-JUL-1997.
XX
PF   14-JUL-1994;   94US-0275076.
XX
PR   04-DEC-1989;   89US-0445302.
PR   15-JAN-1988;   88US-0144357.
PR   28-OCT-1991;   91US-0785865.
PR   14-JUL-1994;   94US-0275076.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Kumar AA,  Mulvihill ER;
XX
DR   WPI; 1997-372063/34.
DR   N-PSDB; AAT89686.
XX
PT   Production of recombinant plasminogen - by co-expression with
PT   plasminogen-processing or -stabilising protein
XX
PS   Example 2; Fig 6A-D; 32pp; English.
XX
CC   This is the protein plasminogen. The encoding cDNA is used in a new
CC   process for the production of plasminogen where a first DNA sequence
CC   encoding plasminogen and at least one additional DNA sequence encoding a
CC   protein that processes or stabilises the plasminogen is introduced into
CC   a eukaryotic host cell. The protein is selected from alpha-1-antitrypsin
CC   (AAT) and its variants and Argserpins. Both the DNA sequences are
CC   operably linked to transcriptional promoter and terminator sequences. The
CC   host cell is cultured under conditions that allow the DNA sequences to be
CC   expressed and the recombinant plasminogen is isolated from the host cell.
CC   Co-expression of plasminogen and the protein gives increased yields of
CC   undegraded plasminogen.
XX
```

```
SQ   Sequence   810 AA;

  Query Match            97.1%;  Score 134;  DB 18;  Length 810;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 44
AAY08685
ID   AAY08685 standard; Protein; 810 AA.
XX
AC   AAY08685;
XX
DT   10-AUG-1999  (first entry)
XX
DE   Human plasminogen protein.
XX
KW   Plasminogen; human; angiostatin; endostatin; gene therapy; vector;
KW   anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
KW   tumour growth; solid tumour; diabetic retinopathy; retina.
XX
OS   Homo sapiens.
XX
PN   WO9926480-A1.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-US24950.
XX
PR   20-NOV-1997;   97US-0975424.
XX
PA   (GENE-) GENETIX PHARM INC.
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Bachelot T,  Leboulch P,  Pawliuk RJ;
XX
DR   WPI; 1999-357696/30.
DR   N-PSDB; AAX77711.
XX
PT   Anti-angiogenic gene therapy vectors
XX
PS   Disclosure; Fig 5; 83pp; English.
XX
CC   This invention describes a novel viral gene therapy vector comprising a
CC   nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC   from human or murine angiostatin, human or murine endostatin and
CC   angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC   sufficiently attenuated for use in human gene therapy. The products of
CC   the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC   ophthalmological activity. The vector is used in gene therapy for
CC   inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC   expresses an anti-angiogenic polypeptide. An additional use comprises
CC   treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC   inhibits angiogenesis in the vicinity of the retina. The vector is
CC   administered to cells ex vivo and then administered to the patient.
XX
SQ   Sequence   810 AA;

  Query Match            97.1%;  Score 134;  DB 20;  Length 810;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 45
AAY02114
ID   AAY02114 standard; Protein; 810 AA.
XX
AC   AAY02114;
XX
DT   16-JUL-1999  (first entry)
XX
DE   SEQ ID 77 of WO9916889.
XX
KW   Angiostatin; endostatin; interferon; thrombospondin;
KW   interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW   anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW   cancer; diabetic retinopathy; macular degeneration; arthritis;
KW   tumor cell production.
XX
OS   Homo sapiens.
XX
PN   WO9916889-A1.
XX
PD   08-APR-1999.
XX
PF   30-SEP-1998;   98WO-US20464.
XX
PR   01-OCT-1997;   97US-0060609.
XX
PA   (SEAR ) SEARLE & CO G D.
XX
PI   Bolanowski MA,  Caparon MH,  Casperson GF,  Gregory SA;
PI   Klein BK,  McKearn JP;
XX
DR   WPI; 1999-255098/21.
XX
PT   New multifunctional proteins useful for treating angiogenic-mediated
PT   diseases
XX
PS   Disclosure; Page 107-109; 121pp; English.
XX
CC   The specification describes multifunctional proteins which comprise
CC   combinations of angiostatin, endostatin, interferon, thrombospondin,
CC   interferon-inducible protein and platelet factor 4, and have
CC   anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC   may exhibit useful properties such as having similar or greater
CC   biological activity when compared to a single factor or by having
CC   improved half-life or decreased adverse side effects, or a combination
CC   of these properties. The proteins can be used for treating an
CC   angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC   degeneration, or arthritis. They can also be used for inhibiting the
CC   production of tumor cells (characteristic of lung, breast, ovarian,
CC   prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC   hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC   growth. The present sequence is used in the course of the invention.
XX
SQ   Sequence   810 AA;

  Query Match            97.1%;  Score 134;  DB 20;  Length 810;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 46
AAY82690
ID   AAY82690 standard; Protein; 810 AA.
XX
AC   AAY82690;
XX
DT   09-AUG-2000  (first entry)
XX
DE   Human plasminogen angiostatin converting enzyme of pH4 (PACE4).
XX
KW   Human; cathepsin D precursor; PACE4; cancer; metastasis; inhibition;
```

KW plasminogen angiostatin converting enzyme of pH4; cytostatic;
KW antidiabetic; ophthalmological; antirheumatic; solid tumour;
KW diabetic retinopathy; rheumatism.
XX
OS Homo sapiens.
XX
PN WO200020570-A1.
XX
PD 13-APR-2000.
XX
PF 29-SEP-1999; 99WO-JP05322.
XX
PR 02-OCT-1998; 98JP-0296095.
XX
PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX
PI Morikawa W, Kaminaka K, Takemoto S, Maeda H, Nozaki C, Miyamoto S;
XX
DR WPI; 2000-303762/26.
XX
PT Enzyme useful for treatment and prevention of tumor metastasis and
PT other diseases involving angiogenesis, splits plasma proteins to give
PT fragments capable of inhibiting cancer metastasis -
XX
PS Disclosure; Fig 7; 55pp; Japanese.
XX
CC The present invention describes an enzyme (I) which splits plasma
CC proteins to give fragments which are capable of inhibiting cancer
CC metastasis. (I) has cytostatic, antidiabetic, ophthalmological and
CC antirheumatic activities. (I) splits plasma proteins to produce
CC fragments which can inhibit metastasis. (I) is useful for the
CC treatment and prevention of diseases in which angiogenesis plays a
CC part including solid tumours, diabetic retinopathy, and rheumatism.
CC The present sequence represents human plasminogen angiostatin converting
CC enzyme of pH4 (PACE4), which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 810 AA;

Query Match 97.1%; Score 134; DB 21; Length 810;
Best Local Similarity 95.7%; Pred. No. 4.9e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 532 RNPDGDVGGPWCYTTNPRKLYDY 554

RESULT 47
AAY53867
ID AAY53867 standard; protein; 810 AA.
XX
AC AAY53867;
XX
DT 13-MAR-2000 (first entry)
XX
DE Amino acid sequence of human plasminogen protein.
XX
KW Human; plasminogen; angiostatin; greenstatin; thrombolytic factor;
KW angiogenesis inhibitory protein; proliferation; angiogenesis; cancer;
KW vascular endothelial cell; ophthalmic disease; glaucoma;
KW diabetic retinopathy; arthritis; psoriasis.
XX
OS Homo sapiens.
XX
PN WO9961464-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-KR00263.
XX
PR 28-MAY-1998; 98KR-0019535.
PR 27-MAY-1999; 99KR-0019144.
XX
PA (GREC) KOREA GREEN CROSS CORP.
XX
PI You WK, So SH, Ahn BC, Lee H, Jung S, Kim Y, Lee JH, Hong Y;
PI Joe YA, Chang S;
XX
DR WPI; 2000-086703/07.
XX
PT Purifying angiogenesis inhibitors produced as recombinant proteins in
PT Escherichia coli, useful as anticancer agents and for treating ocular
PT diseases -
XX
PS Disclosure; Page 41-45; 55pp; English.
XX
CC The present sequence represents the human plasminogen protein. Fragments
CC of the plasminogen protein, specifically angiostatin (comprising amino
CC acids 99-467) and greenstatin (comprising amino acids 101-354) are used
CC as thrombolytic factors and angiogenesis inhibitory proteins. Angiostatin
CC contains the kringle 1-4 region of plasminogen, and greenstatin contains
CC the kringle 1-3 region of plasminogen. As both proteins contain a high
CC number of disulphide bonds, they are difficult to purify. The
CC specification describes a method for the purification of such
CC angiogenesis inhibitory proteins. The method comprises solubilising
CC the proteins, produced as inclusion bodies in Escherichia coli and
CC refolding the solubilised fraction in buffer containing urea and
CC glutathione. The angiogenesis inhibitory proteins specifically inhibit
CC proliferation of vascular endothelial cells, but not that of
CC non-endothelial cancers or normal cells. The angiogenesis inhibitory
CC proteins are used to suppress angiogenesis, specifically for treating
CC cancers (e.g. of lung, skin and brain) or ophthalmic diseases (e.g.
CC glaucoma and diabetic retinopathy), but also arthritis and psoriasis.
XX
SQ Sequence 810 AA;

Query Match 97.1%; Score 134; DB 21; Length 810;
Best Local Similarity 95.7%; Pred. No. 4.9e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
||||||||||| ||||||||||
Db 532 RNPDGDVGGPWCYTTNPRKLYDY 554

RESULT 48
AAR12933
ID AAR12933 standard; Protein; 811 AA.
XX
AC AAR12933;
XX
DT 23-SEP-1991 (first entry)
XX
DE Plasminogen mutein X1 with factor Xa cleavage site.
XX
KW protease; fibrinolysis; blood clotting; thrombosis.
XX
FH Key Location/Qualifiers
FT Cleavage-site 578..581
FT /label= Factor Xa cleavage site
XX
PN WO9109118-A.
XX
PD 27-JUN-1991.
XX
PF 07-DEC-1990; 90WO-GO01912.
XX
PR 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90WO-GB01911.
XX
PA (BRBI-) BRIT BIO-TECHN LTD.
XX
PI Dawson KM, Edwards RM, Forman JM;
XX
DR WPI; 1991-208145/28.

```
DR   N-PSDB; AAQ12542.
XX
PT   Activatable fibrinolytic and antithrombic proteins - activated by
PT   e.g. factor Xa, thrombin or activated protein C
XX
PS   Claim 7; Fig 2 and Fig 4; 73pp; English.
XX
CC   This protein is a plasminogen mutant cleavable by Factor Xa.
CC   Activation is localised to the thrombus because cleavage to plasmin
CC   is by an enzyme of the blood clotting pathway. Compositions
CC   comprising the mutant plasminogen are used for treatment or
CC   prevention of thrombosis, etc.
CC   See AAQ12543-Q12558.
XX
SQ   Sequence   811 AA;

  Query Match            97.1%;  Score 134;  DB 12;  Length 811;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||||||||||  ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 49
AAR12939
ID   AAR12939 standard; Protein; 811 AA.
XX
AC   AAR12939;
XX
DT   23-SEP-1991  (first entry)
XX
DE   Plasminogen mutein T2 with thrombin cleavage site.
XX
KW   protease; fibrinolysis; blood clotting; thrombosis.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   578..581
FT                   /label= Thrombin cleavage site
XX
PN   WO9109118-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1990;   90WO-GO01912.
XX
PR   07-DEC-1989;   89GB-0027722.
PR   07-DEC-1990;   90WO-GB01911.
XX
PA   (BRBI-) BRIT BIO-TECHN LTD.
XX
PI   Dawson KM,  Edwards RM,  Forman JM;
XX
DR   WPI; 1991-208145/28.
DR   N-PSDB; AAQ12548.
XX
PT   Activatable fibrinolytic and antithrombic proteins - activated by
PT   e.g. factor Xa, thrombin or activated protein C
XX
PS   Claim 7; Fig 2 and Fig 5; 73pp; English.
XX
CC   This protein is a plasminogen mutant cleavable by thrombin.
CC   Activation is localised to the thrombus because cleavage to plasmin
CC   is by an enzyme of the blood clotting pathway. Compositions
CC   comprising the mutant plasminogen are used for treatment or
CC   prevention of thrombosis, etc.
CC   See AAQ12542-Q12558.
XX
SQ   Sequence   811 AA;

  Query Match            97.1%;  Score 134;  DB 12;  Length 811;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||||||||||  ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 50
AAR12943
ID   AAR12943 standard; Protein; 811 AA.
XX
AC   AAR12943;
XX
DT   23-SEP-1991  (first entry)
XX
DE   Plasminogen mutein T13 with thrombin cleavage site.
XX
KW   protease; fibrinolysis; blood clotting; thrombosis.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   579..583
FT                   /note= "recognised by  thrombin"
XX
PN   WO9109118-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1990;   90WO-GB01912.
XX
PR   07-DEC-1989;   89GB-0027722.
PR   07-DEC-1990;   90WO-GB01911.
XX
PA   (BRBI-) BRIT BIO-TECHN LTD.
XX
PI   Dawson KM,  Edwards RM,  Forman JM;
XX
DR   WPI; 1991-208145/28.
DR   N-PSDB; AAQ12552.
XX
PT   Activatable fibrinolytic and antithrombic proteins - activated by
PT   e.g. factor Xa, thrombin or activated protein C
XX
PS   Claim 9; Fig 2 and Fig 5; 73pp; English.
XX
CC   This protein is a plasminogen mutant cleavable by thrombin.
CC   Activation is localised to the thrombus because cleavage to plasmin
CC   is by an enzyme of the blood clotting pathway. Compositions
CC   comprising the mutant plasminogen are used for treatment or
CC   prevention of thrombosis, etc.
CC   See AAQ12542-Q12558.
XX
SQ   Sequence   811 AA;

  Query Match            97.1%;  Score 134;  DB 12;  Length 811;
  Best Local Similarity  95.7%;  Pred. No. 4.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||||||||||  ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


Search completed: November  8, 2002, 09:33:00
Job time : 33 secs
```

(9)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:31:20 ; Search time 15 Seconds
(without alignments)
147.337 Million cell updates/sec

Title: US-09-657-431-9
Perfect score: 138
Sequence: 1 RNPDGDVGGPWAYTTNPRKLYDY 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	97.1	810	1	PLHU	plasmin (EC 3.4.21
2	130	94.2	810	2	B30848	plasmin (EC 3.4.21
3	128	92.8	812	1	PLMS	plasmin (EC 3.4.21
4	124	89.9	460	2	B61545	plasmin (EC 3.4.21
5	120	87.0	810	2	I46260	plasmin (EC 3.4.21
6	117	84.8	4548	1	S00657	apoprotein(a) (EC
7	114	82.6	455	2	A61545	plasmin (EC 3.4.21
8	113	81.9	790	1	PLPG	plasmin (EC 3.4.21
9	111	80.4	812	1	PLBO	plasmin (EC 3.4.21
10	94	68.1	716	1	JC5061	macrophage-stimula
11	92	66.7	711	1	A47136	macrophage-stimula
12	90	65.2	716	1	A40332	macrophage-stimula
13	82	59.4	728	1	A35644	hepatocyte growth
14	81	58.7	728	1	JH0579	hepatocyte growth
15	81	58.7	728	1	A60185	hepatocyte growth
16	80	58.0	625	1	TBBO	thrombin (EC 3.4.2
17	78	56.5	710	1	I51283	hepatocyte growth
18	77	55.8	169	2	A40522	plasmin (EC 3.4.21
19	75	54.3	411	2	I51285	hepatocyte growth
20	73	52.9	336	2	S33879	plasmin precursor
21	73	52.9	617	2	S10511	thrombin (EC 3.4.2
22	72.5	52.5	2869	2	T18518	apolipoprotein(a)
23	71	51.4	618	2	A35827	thrombin (EC 3.4.2
24	71	51.4	622	1	TBHU	thrombin (EC 3.4.2
25	68.5	49.6	123	2	C61545	plasmin (EC 3.4.21
26	66.5	48.2	89	2	A60140	plasmin (EC 3.4.21
27	65.5	47.5	1420	2	A32869	apolipoprotein(a)
28	62.5	45.3	120	2	E61545	plasmin (EC 3.4.21
29	60	43.5	435	2	E70711	hypothetical prote
30	59	42.8	943	2	B45082	neurotrophic recep
31	56	40.6	559	1	A35029	t-plasminogen acti
32	56	40.6	559	1	A29941	t-plasminogen acti
33	53.5	38.8	153	2	T37219	probable lipoprote
34	52	37.7	323	2	T25094	hypothetical prote
35	52	37.7	562	1	UKHUT	t-plasminogen acti
36	51	37.0	194	2	T10851	y4hQ protein - Rhi
37	51	37.0	552	2	S36786	carboxylesterase (
38	51	37.0	564	1	S36787	carboxylesterase (
39	50.5	36.6	394	2	JS0600	t-plasminogen acti
40	49	35.5	297	2	E83194	polyamine transpor
41	49	35.5	539	2	G70520	probable csp prote
42	49	35.5	615	1	KFHU12	coagulation factor
43	49	35.5	946	1	A47299	ror-related recept
44	48.5	35.1	961	2	T49228	plasma membrane H+
45	48	34.8	268	2	B64318	hypothetical prote
46	48	34.8	347	2	T32768	hypothetical prote
47	48	34.8	357	2	T32881	hypothetical prote
48	48	34.8	410	2	AC3603	2-oxoisovalerate d
49	48	34.8	508	2	F70662	probable plcC prot
50	47	34.1	461	2	H84099	cell wall-binding

ALIGNMENTS

RESULT 1
PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N;Alternate names: plasminogen precursor [misnomer]
N;Contains: angiostatin; microplasmin; plasminogen
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C;Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the f
A;Reference number: A35229; MUID:90202879
A;Accession: A35229
A;Molecule type: DNA
A;Residues: 1-810 <PET>
A;Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A;Experimental source: leukocyte; lung fibroblast
R;Malgaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A;Title: Definition of the transcription initiation site of human plasminogen gene in
A;Reference number: I52242; MUID:91097523
A;Accession: I52242
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <MAL1>
A;Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613
R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human
A;Reference number: A26646; MUID:87162490
A;Accession: A26646
A;Molecule type: mRNA
A;Residues: 1-471,'D',473-810 <FOR>
A;Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
A;Experimental source: liver
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human a
A;Reference number: I45961; MUID:85023311
A;Accession: I62738
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 292-471,'D',473-810 <MAL2>
A;Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A;Accession: I84609
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 367-419 <MAL3>
A;Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A;Reference number: S03735; MUID:81212097
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71,'E',73-76 <BRU>
R;Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: protein
A;Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R;Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A;Title: Primary structure of the B-chain of human plasmin.
A;Reference number: A04627; MUID:77225245
A;Accession: A04627
A;Molecule type: protein
A;Residues: 581-810 <WI1>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla

A;Reference number: A04625; MUID:75093329
A;Accession: A04625
A;Molecule type: protein
A;Residues: 20-50,'Q',51-71,'E',73-85,87-100 <WI2>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A;Reference number: A04626; MUID:76043692
A;Accession: A04626
A;Molecule type: protein
A;Residues: 483-507,'E',509-604 <WI3>
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A;Title: The primary structure of human plasminogen. II. The histidine loop of human pla
A;Reference number: A92125; MUID:73149248
A;Contents: annotation; active site
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A;Reference number: A92048; MUID:69234739
A;Contents: annotation; active site
R;Trexler, M.; Vali, Z.; Patthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A;Reference number: A92382; MUID:82213905
A;Contents: annotation; omega-aminocarboxylic acid binding sites
R;Vali, Z.; Patthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A;Reference number: A92458; MUID:85054794
A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A;Title: Kringle domains of human angiostatin. Characterization of the anti-proliferativ
A;Reference number: A58811; MUID:97067211
A;Contents: annotation
R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A;Title: Generation of an angiostatin-like fragment from plasminogen by stromelysin-1 (M
A;Reference number: A58812; MUID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR
A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PMK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.
A;Reference number: A58819; MUID:92031502
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plas
A;Reference number: A58818; MUID:92031503
A;Contents: annotation
R;de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook,
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at
A;Reference number: A39483; MUID:92118803
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
R;Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmi
A;Reference number: A58817; MUID:94237158
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH
d PIR:FGHUGB).
C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately
rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibit
C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial c
C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiostatin.
ting solid tumors.
C;Genetics:
A;Gene: GDB:PLG
A;Cross-references: GDB:119498; OMIM:173350
A;Map position: 6q26-6q27
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1;
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol

```
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;20-96/Domain: activation peptide #status experimental <APT>
F;79-466/Product: angiostatin #status experimental <AST>
F;97-580,581-810/Product: plasmin #status experimental <MAT>
F;97-580/Domain: plasmin chain A #status experimental <CHA>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;550-580,581-810/Product: microplasmin #status experimental <MMT>

  Query Match             97.1%;  Score 134;  DB 1;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 1.6e-11;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||||
Db    532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 2
B30848
plasmin (EC 3.4.21.7) precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C;Accession: B32869; B30848
R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A;Reference number: A32869; MUID:89174660
A;Accession: B32869
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-810 <TOM>
A;Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-9/Domain: signal sequence #status predicted <SIG>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;581-803/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
 bonds: #status predicted
F;622,665,760/Active site: His, Asp, Ser #status predicted

  Query Match             94.2%;  Score 130;  DB 2;  Length 810;
  Best Local Similarity   91.3%;  Pred. No. 6.3e-11;
  Matches   21;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||||||| ||||||||:||
Db    532 RNPDGDVGGPWCYTTNPRKLFDY 554


RESULT 3
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiostatin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of t
A;Reference number: A38514; MUID:91184812
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DEG>
A;Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LI2>
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, both plas
mediately after dissociation from the clot. In the presence of the inhibitor, the act
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiostatin.
eful in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiostatin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
 bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

  Query Match             92.8%;  Score 128;  DB 1;  Length 812;
  Best Local Similarity   91.3%;  Pred. No. 1.2e-10;
  Matches   21;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||| ||| ||||||||||
Db    532 RNPDGDVNGPWCYTTNPRKLYDY 554


RESULT 4
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37;38-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
```

```
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; MUID:93149995
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z
F;1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
F;1-37/Domain: activation peptide (fragment) #status experimental <APT>
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F;41-118/Domain: kringle homology <KR4>
F;118-460/Product: miniplasminogen #status experimental <MIN>
F;132-211/Domain: kringle homology <KR5>
F;226-460/Domain: plasmin chain B #status experimental <BCH>
F;231-453/Domain: trypsin homology <TRY>
F;272,315,410/Active site: His, Asp, Ser #status predicted

  Query Match             89.9%;  Score 124;  DB 2;  Length 460;
  Best Local Similarity   87.0%;  Pred. No. 2.6e-10;
  Matches   20;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| ||||||||:||
Db     183 RNPDGDVNGPWCYTTNPRKLFDY 205


RESULT 5
I46260
plasmin (EC 3.4.21.7) precursor - western European hedgehog
C;Species: Erinaceus europaeus (western European hedgehog)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46260
R;Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A;Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprot
A;Reference number: I46259; MUID:96025778
A;Accession: I46260
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-810 <LAW>
A;Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: hydrolase; serine proteinase
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;379-456/Domain: kringle homology <KR4>
F;482-561/Domain: kringle homology <KR5>
F;582-803/Domain: trypsin homology <TRY>

  Query Match             87.0%;  Score 120;  DB 2;  Length 810;
  Best Local Similarity   82.6%;  Pred. No. 1.8e-09;
  Matches   19;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  ||| ||||||||:||
Db     533 RNPDGDANGPWCYTTNPRKLFDY 555


RESULT 6
S00657
apoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N;Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-2000
C;Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R;McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scan
Nature 330, 132-137, 1987
A;Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A;Reference number: S00657; MUID:88039109
A;Accession: S00657
A;Molecule type: mRNA
A;Residues: 1-4548 <MCL>
A;Cross-references: GB:X06290; EMBL:X06696; NID:g28619; PIDN:CAA29618.1; PID:g28620
R;Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A;Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous
A;Reference number: A28017; MUID:87204109
A;Accession: A28017
A;Molecule type: protein
A;Residues: 20-21,'P',23-34;177-179,'N',181-186,'T',188-196,'DKG',200;292-314,'W',316
X',4396-4401 <EAT>
R;Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A;Title: 5' control regions of the apolipoprotein(a) gene and members of the related
A;Reference number: A47277; MUID:93165698
A;Accession: A47277
A;Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RES>
A;Cross-references: GB:L07899; NID:g967973; PID:g967974
R;Malgaretti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sa
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A;Title: Characterization by yeast artificial chromosome cloning of the linked apolip
A;Reference number: A47233; MUID:93087573
A;Accession: I60906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE2>
A;Cross-references: GB:M90078; NID:g178786; PIDN:AAA35547.1; PID:g553188
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: A47233
A;Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE5>
A;Cross-references: GB:M90079; NID:g178784; PIDN:AAA35546.1; PID:g553187
R;Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A;Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated
A;Reference number: I52415; MUID:92207924
A;Accession: I52415
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE3>
A;Cross-references: GB:M86877; NID:g178780; PIDN:AAB49909.1; PID:g553185
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: I65286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE4>
A;Cross-references: GB:M86878; NID:g178782; PIDN:AAA51749.1; PID:g553186
C;Genetics:
A;Gene: GDB:LPA
A;Cross-references: GDB:120699; OMIM:152200
A;Map position: 6q26-6q27
A;Note: several genes closely linked on chromosome 6 are identical in the first codin
rs of kringle repeats
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F;28-105/Domain: kringle homology <KR1>
F;142-219/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;370-447/Domain: kringle homology <KR4>
F;484-561/Domain: kringle homology <KR5>
F;598-675/Domain: kringle homology <KR6>
F;712-789/Domain: kringle homology <KR7>
F;826-903/Domain: kringle homology <KR8>
F;940-1017/Domain: kringle homology <KR9>
F;1054-1131/Domain: kringle homology <KR10>
F;1168-1245/Domain: kringle homology <KR11>
F;1282-1359/Domain: kringle homology <KR12>
F;1396-1473/Domain: kringle homology <KR13>
```

F;1510-1587/Domain: kringle homology <KR14>
F;1624-1701/Domain: kringle homology <KR15>
F;1738-1815/Domain: kringle homology <KR16>
F;1852-1929/Domain: kringle homology <KR17>
F;1966-2043/Domain: kringle homology <KR18>
F;2080-2157/Domain: kringle homology <KR19>
F;2194-2271/Domain: kringle homology <KR20>
F;2308-2385/Domain: kringle homology <KR21>
F;2422-2499/Domain: kringle homology <KR22>
F;2536-2613/Domain: kringle homology <KR23>
F;2650-2727/Domain: kringle homology <KR24>
F;2764-2841/Domain: kringle homology <KR25>
F;2878-2955/Domain: kringle homology <KR26>
F;2992-3069/Domain: kringle homology <KR27>
F;3106-3183/Domain: kringle homology <KR28>
F;3220-3297/Domain: kringle homology <KR29>
F;3334-3411/Domain: kringle homology <KR30>
F;3448-3525/Domain: kringle homology <KR31>
F;3562-3639/Domain: kringle homology <KR32>
F;3676-3753/Domain: kringle homology <KR33>
F;3782-3859/Domain: kringle homology <KR34>
F;3896-3973/Domain: kringle homology <KR35>
F;4010-4087/Domain: kringle homology <KR36>
F;4124-4201/Domain: kringle homology <KR37>
F;4228-4307/Domain: kringle homology <KR38>
F;4328-4541/Domain: trypsin homology <TRY>

Query Match 84.8%; Score 117; DB 1; Length 4548;
Best Local Similarity 78.3%; Pred. No. 3.3e-08;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||||: ||| || |||||:||
Db    4279 RNPDGDINGPWCYTMNPRKLFDY 4301
```

RESULT 7
A61545
plasmin (EC 3.4.21.7) precursor - horse (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Equus caballus (domestic horse)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: A61545; S17527
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: A61545
A;Molecule type: protein
A;Residues: 1-33;34-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A;Title: Complete amino acid sequence of equine miniplasminogen.
A;Reference number: S17527; MUID:92052077
A;Accession: S17527
A;Molecule type: protein
A;Residues: 118-455 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z
F;1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F;1-33/Domain: activation peptide (fragment) #status experimental <APT>
F;34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F;37-114/Domain: kringle homology <KR4>
F;118-455/Product: miniplasminogen #status experimental <MIN>
F;126-205/Domain: kringle homology <KR5>
F;226-455/Domain: plasmin chain B #status experimental <BCH>
F;226-448/Domain: trypsin homology <TRY>
F;267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 82.6%; Score 114; DB 2; Length 455;
Best Local Similarity 78.3%; Pred. No. 7.4e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| || ||:||:||
Db     177 RNPDGDVNGPWCYTMNPQKLFDY 199
```

RESULT 8
PLPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: S03733; S03737; A25834
R;Schaller, J.; Marti, T.; Roesselet, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A;Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the
A;Reference number: S03733
A;Accession: S03733
A;Molecule type: protein
A;Residues: 1-560 <SCH>
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg,
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of huma
A;Reference number: S03735; MUID:81212097
A;Accession: S03737
A;Molecule type: protein
A;Residues: 1-57 <BRU>
R;Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A;Title: Determination of the complete amino-acid sequence of porcine miniplasminogen
A;Reference number: A25834; MUID:85203907
A;Accession: A25834
A;Molecule type: protein
A;Residues: 450-790 <MAR>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine pr
F;1-790/Product: plasminogen #status predicted <PRO>
F;1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F;1-77/Domain: activation peptide #status predicted <APT>
F;78-560/Product: plasmin chain A #status predicted <ACH>
F;84-162/Domain: kringle homology <KR1>
F;166-243/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;358-435/Domain: kringle homology <KR4>
F;450-790/Product: miniplasminogen #status experimental <MIN>
F;461-540/Domain: kringle homology <KR5>
F;561-790/Product: plasmin chain B #status experimental <BCH>
F;561-783/Domain: trypsin homology <TRY>
F;30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,
bonds: #status predicted
F;602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 81.9%; Score 113; DB 1; Length 790;
Best Local Similarity 78.3%; Pred. No. 1.9e-08;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||| ||| |||||:||:||
Db     512 RNPDGDDNGPWCYTTNPQKLFDY 534
```

RESULT 9
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N;Alternate names: plasminogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C;Accession: S45046; A25835; I45961; S03736

R;Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A;Description: Cloning and characterizatin of the bovine plasminogen cDNA.
A;Reference number: S45046
A;Accession: S45046
A;Molecule type: mRNA
A;Residues: 1-812 <BER>
A;Cross-references: EMBL:X79402; NID:g494962; PIDN:CAA55939.1; PID:g494963
A;Experimental source: liver
A;Note: it is uncertain whether Met-1 or Met-8 is the initiator
R;Schaller, J.; Moser, P.W.; Dannegger-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Rick
Eur. J. Biochem. 149, 267-278, 1985
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasm
A;Reference number: A25835; MUID:85203906
A;Accession: A25835
A;Molecule type: protein
A;Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A;Reference number: I45961; MUID:85023311
A;Accession: I45961
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 706-743,'R',745-812 <MAL>
A;Cross-references: GB:K02935; NID:g163551; PIDN:AAA30714.1; PID:g163552
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A;Reference number: S03735; MUID:81212097
A;Accession: S03736
A;Molecule type: protein
A;Residues: 27-83 <BRU>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma;
F;1-26/Domain: signal sequence #status predicted <SIG>
F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F;27-812/Product: plasminogen #status experimental <PRO>
F;27-103/Domain: activation peptide #status experimental <APT>
F;104-583,584-812/Product: plasmin #status experimental <MAT>
F;104-583/Domain: plasmin chain A #status experimental <ACH>
F;110-188/Domain: kringle homology <KR1>
F;192-269/Domain: kringle homology <KR2>
F;282-359/Domain: kringle homology <KR3>
F;384-461/Domain: kringle homology <KR4>
F;485-564/Domain: kringle homology <KR5>
F;584-812/Domain: plasmin chain B #status experimental <BCH>
F;584-805/Domain: trypsin homology <TRY>
F;56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,33
bonds: #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 80.4%; Score 111; DB 1; Length 812;
Best Local Similarity 78.3%; Pred. No. 3.8e-08;
Matches 18; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| || |||| :||
Db     536 RNPDGDVNGPWCYTMNPRKPFDY 558
```

RESULT 10
JC5061
macrophage-stimulating protein 1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C;Accession: JC5061

R;Ohshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A;Title: Molecular cloning of rat macrophage-stimulating protein and its involvement
A;Reference number: JC5061; MUID:97011126
A;Accession: JC5061
A;Molecule type: mRNA
A;Residues: 1-716 <OHS>
A;Cross-references: EMBL:X95096; NID:g1669718; PIDN:CAA64473.1; PID:g1669719
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F;32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F;110-186/Domain: kringle homology <KRI1>
F;191-268/Domain: kringle homology <KRI2>
F;292-370/Domain: kringle homology <KRI3>
F;379-457/Domain: kringle homology <KRI4>
F;489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F;489-709/Domain: trypsin homology <TRY>
F;72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.1%; Score 94; DB 1; Length 716;
Best Local Similarity 65.2%; Pred. No. 1e-05;
Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  ||| || :|  |:||
Db     429 RNPDGDSHGPWCYTLDPETLFDY 451
```

RESULT 11
A47136
macrophage-stimulating protein 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C;Accession: A40331; B40331; A47136; A61395
R;Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A;Title: Characterization of the DNF15S2 locus on human chromosome 3: identification
A;Reference number: A40331; MUID:92002016
A;Accession: A40331
A;Molecule type: DNA
A;Residues: 1-711 <HA1>
A;Cross-references: GB:M74179
A;Accession: B40331
A;Molecule type: mRNA
A;Residues: 1-711 <HA2>
A;Cross-references: GB:M74178; NID:g183976; PIDN:AAA50165.1; PID:g183977
R;Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A;Title: Cloning, sequencing, and expression of human macrophage stimulating protein
A;Reference number: A47136; MUID:93340141
A;Accession: A47136
A;Molecule type: mRNA
A;Residues: 1-12,'C',14-622,'F',624-711 <YOS>
A;Cross-references: GB:L11924; NID:g398037; PIDN:AAA59872.1; PID:g398038
A;Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequen
R;Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A;Title: Macrophage stimulating protein: purification, partial amino acid sequence, a
A;Reference number: A61395; MUID:91217635
A;Accession: A61395
A;Molecule type: protein
A;Residues: 230-247;288-291,'E',293-295,'X',297-301,'X',303,'E',305,'EX',308-310;326-
A;Experimental source: plasma
C;Genetics:
A;Gene: GDB:MST1; D3F15S2; DNF15S2; HGFL
A;Cross-references: GDB:128833; OMIM:142408
A;Map position: 3p21-3p21.3
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle; plasma

```
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F;19-483/Domain: alpha chain #status predicted <ACH>
F;110-186/Domain: kringle homology <KR1>
F;191-268/Domain: kringle homology <KR2>
F;283-361/Domain: kringle homology <KR3>
F;370-448/Domain: kringle homology <KR4>
F;484-711/Domain: beta chain #status predicted <BCH>
F;484-704/Domain: trypsin homology <TRY>
F;56-78,60-66,110-186,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355,37
F;72,296,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             66.7%;  Score 92;  DB 1;  Length 711;
  Best Local Similarity   65.2%;  Pred. No. 2e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  ||| || :||  :||
Db     420 RNPDGDSHGPWCYTMDPRTPFDY 442


RESULT 12
A40332
macrophage-stimulating protein 1 precursor - mouse
N;Alternate names: hepatocyte growth factor-like protein
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C;Accession: A40332; B40332
R;Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
Biochemistry 30, 9781-9791, 1991
A;Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth fact
A;Reference number: A40332; MUID:92002017
A;Accession: A40332
A;Molecule type: DNA
A;Residues: 1-716 <DEG>
A;Cross-references: GB:M74180; NID:g193831; PIDN:AAA50166.1; PID:g193832
A;Accession: B40332
A;Molecule type: mRNA
A;Residues: 1-18,'P',20-716 <DEG2>
A;Cross-references: GB:M74181; NID:g193833; PIDN:AAA50167.1; PID:g193834
C;Genetics:
A;Introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1; 47
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle
F;1-31/Domain: signal sequence #status predicted <SIG>
F;19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
F;19-483/Domain: alpha chain #status experimental <ACH>
F;110-186/Domain: kringle homology <KR1>
F;191-268/Domain: kringle homology <KR2>
F;292-370/Domain: kringle homology <KR3>
F;379-457/Domain: kringle homology <KR4>
F;484-711/Domain: beta chain #status experimental <BCH>
F;489-709/Domain: trypsin homology <TRY>
F;72,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             65.2%;  Score 90;  DB 1;  Length 716;
  Best Local Similarity   65.2%;  Pred. No. 3.9e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  ||| || :|  |:||
Db     429 RNPDGDSHGPWCYTLDPDILFDY 451


RESULT 13
A35644
hepatocyte growth factor precursor - rat
N;Alternate names: hepapoietin A; scatter factor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
C;Accession: A35644; S13211
R;Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Nakamur
Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
A;Title: Deduced primary structure of rat hepatocyte growth factor and expression of
A;Reference number: A35644; MUID:90222197
A;Accession: A35644
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-728 <TAS>
A;Cross-references: GB:D90102; GB:M32987; NID:g220766; PIDN:BAA14133.1; PID:g220767
A;Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 4
R;Okajima, A.; Miyazawa, K.; Kitamura, N.
Eur. J. Biochem. 193, 375-381, 1990
A;Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA
A;Reference number: S13211; MUID:91031482
A;Accession: S13211
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-728 <OKA>
A;Cross-references: EMBL:X54400; NID:g56353; PIDN:CAA38266.1; PID:g4539554
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
F;1-32/Domain: signal sequence #status predicted <SIG>
F;56-495/Product: hepatocyte growth factor #status predicted <MAT>
F;56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;129-207/Domain: kringle homology <KR1>
F;212-289/Domain: kringle homology <KR2>
F;306-384/Domain: kringle homology <KR3>
F;392-470/Domain: kringle homology <KR4>
F;496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;496-719/Domain: trypsin homology <TRY>
F;33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;295,403,569,656/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;488-607/Disulfide bonds: #status predicted

  Query Match             59.4%;  Score 82;  DB 1;  Length 728;
  Best Local Similarity   60.9%;  Pred. No. 0.0006;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || ||   :||
Db     442 RNPDDDAHGPWCYTGNPLVPWDY 464


RESULT 14
JH0579
hepatocyte growth factor precursor [validated] - human
N;Alternate names: hepapoietin A; scatter factor
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C;Accession: JH0579; JU0333; A41140; B36677; A36677; A33512; A39006; PH0114; A37796;
R;Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A;Title: Organization of the human hepatocyte growth factor-encoding gene.
A;Reference number: JH0579; MUID:91340155
A;Accession: JH0579
A;Molecule type: DNA
A;Residues: 1-728 <SEK>
A;Cross-references: DDBJ:D90318
A;Note: the authors translated the codon GAA for residue 662 as Gly
R;Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
submitted to JIPID, March 1991
A;Description: Organization of the human hepatocyte growth factor-encoding gene.
A;Reference number: JU0333
A;Accession: JU0333
A;Molecule type: DNA
A;Residues: 1-481,'RT',484-728 <SE2>
R;Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder,
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A;Title: Evidence for the identity of human scatter factor and human hepatocyte growt
A;Reference number: A41140; MUID:91334393
```

A;Accession: A41140
A;Molecule type: mRNA
A;Residues: 1-728 <WEI>
A;Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936
R;Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A;Title: Isolation and expression of cDNA for different forms of hepatocyte growth factc
A;Reference number: A36677; MUID:91025062
A;Accession: B36677
A;Molecule type: mRNA
A;Residues: 1-728 <SE3>
A;Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032
A;Accession: A36677
A;Molecule type: mRNA
A;Residues: 1-161,167-728 <SE4>
A;Cross-references: EMBL:X16323
A;Experimental source: leukocyte
R;Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakaya
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A;Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac
A;Reference number: A33512; MUID:89392017
A;Accession: A33512
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-728 <MIY>
A;Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306846
R;Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hir
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A;Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocy
A;Reference number: A39006; MUID:91110540
A;Accession: A39006
A;Molecule type: mRNA
A;Residues: 1-161,167-728 <RUB>
A;Cross-references: GB:M55379
A;Experimental source: embryonic lung
R;Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A;Title: Identification of the N-terminal residue of the heavy chain of both native and
A;Reference number: PH0114; MUID:91207365
A;Accession: PH0114
A;Molecule type: protein
A;Residues: 32-43;53-58 <YOS>
A;Experimental source: plasma
R;Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A;Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
A;Reference number: A37796; MUID:91035621
A;Accession: A37796
A;Molecule type: protein
A;Residues: 86-91;329-344;356-363,'XX',366-370;425-434;442-447,'X',449-450;543-546,'X',5
R;Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tashi
Nature 342, 440-443, 1989
A;Title: Molecular cloning and expression of human hepatocyte growth factor.
A;Reference number: S06794; MUID:90066676
A;Accession: S06794
A;Molecule type: mRNA
A;Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'M',301-316,'A',318-335,'K',337-386,'
A;Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
A;Experimental source: liver
A;Note: the authors translated the codon CAG for residue 727 as Glu
A;Note: part of this sequence, including the amino end of both the alpha and beta chains
R;Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A;Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
A;Reference number: I59214; MUID:93087571
A;Accession: I59214
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-288,'ET' <HAR>
A;Cross-references: GB:L02931; NID:g184033; PIDN:AAA52649.1; PID:g184034
R;Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A;Title: An alternatively processed mRNA generated from human hepatocyte growth factor g

A;Reference number: S15443; MUID:91200041
A;Accession: S15443
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-288,'ET' <MIY2>
A;Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084
R;Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A;Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: clon
A;Reference number: I52253; MUID:92062058
A;Accession: I52253
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 161-166 <SHI>
A;Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997
C;Genetics:
A;Gene: GDB:HGF
A;Cross-references: GDB:127524; OMIM:142409
A;Map position: 7q21.1-7q21.1
A;Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1;
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-494,495-728/Product: hepatocyte growth factor #status experimental <MAT>
F;32-494/Domain: alpha chain #status experimental <ACH>
F;128-206/Domain: kringle homology <KR1>
F;211-288/Domain: kringle homology <KR2>
F;305-383/Domain: kringle homology <KR3>
F;391-469/Domain: kringle homology <KR4>
F;495-728/Domain: beta chain #status experimental <BCH>
F;495-716/Domain: trypsin homology <TRY>
F;32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;294,402,566,653/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;487-604/Disulfide bonds: #status predicted

Query Match 58.7%; Score 81; DB 1; Length 728;
Best Local Similarity 60.9%; Pred. No. 0.00084;
Matches 14; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```
Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||| |  ||| || ||   :||
Db    441 RNPDDDAHGPWCYTGNPLIPWDY 463
```

RESULT 15
A60185
hepatocyte growth factor precursor - mouse
N;Alternate names: hepapoietin A; scatter factor
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1993 #sequence_revision 26-May-1994 #text_change 16-Jun-2000
C;Accession: JC2117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; JU0231
R;Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.
Biochem. Biophys. Res. Commun. 199, 772-779, 1994
A;Title: Identification of mouse mammary fibroblast-derived mammary growth factor as
A;Reference number: JC2117; MUID:94183257
A;Accession: JC2117
A;Molecule type: mRNA
A;Residues: 1-728 <SAS2>
A;Cross-references: GB:D10212; NID:g220435; PIDN:BAA01064.1; PID:g220436
A;Experimental source: fibroblast, COS-1 cell
A;Note: submitted to JIPID, May 1993
A;Accession: PC2064
A;Molecule type: protein
A;Residues: 496-504 <SA2>
R;Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.
Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990
A;Title: Purified scatter factor stimulates epithelial and vascular endothelial cell
A;Reference number: A60185; MUID:90377927
A;Accession: A60185

A;Molecule type: protein
A;Residues: 'X',184-188,'KX',191-192,'X',194,'XX',197;357-364,'XX',367;375-377,'E',379,'
R;Liu, Y.; Michalopoulos, G.K.; Zarnegar, R.
Biochim. Biophys. Acta 1216, 299-303, 1993
A;Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
A;Reference number: S43416; MUID:94060105
A;Accession: S43416
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-728 <LIU>
A;Cross-references: EMBL:X72307
R;Liu, Y.
submitted to the EMBL Data Library, May 1993
A;Reference number: S45521
A;Accession: S45521
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-563,'H',565-728 <LI2>
A;Cross-references: EMBL:X72307
R;Coffer, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
Biochem. J. 278, 35-41, 1991
A;Title: Purification and characterization of biologically active scatter factor from ra
A;Reference number: S17173; MUID:91354223
A;Accession: S17173
A;Molecule type: protein
A;Residues: 496-517,'T',519 <COF>
R;Gherardi, E.; Stoker, M.
Nature 346, 228, 1990
A;Title: Hepatocytes and scatter factor.
A;Reference number: S10966; MUID:90326152
A;Accession: S10966
A;Status: preliminary
A;Molecule type: protein
A;Residues: 496-507,'X',509-512,'L',514-516,'X',518-519 <NAT>
R;Plaschke-Schlutter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
J. Biol. Chem. 270, 830-836, 1995
A;Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
A;Reference number: I48758; MUID:95122532
A;Accession: I48758
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-30 <RES>
A;Cross-references: EMBL:X81630; NID:g673451; PIDN:CAA57286.1; PID:g673452
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
F;1-32/Domain: signal sequence #status predicted <SIG>
F;56-495,496-728/Product: hepatocyte growth factor #status predicted <MAT>
F;56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;129-207/Domain: kringle homology <KR1>
F;212-289/Domain: kringle homology <KR2>
F;306-384/Domain: kringle homology <KR3>
F;392-470/Domain: kringle homology <KR4>
F;496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;496-719/Domain: trypsin homology <TRY>
F;33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;295,403,569,656/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;488-607/Disulfide bonds: #status predicted

Query Match 58.7%; Score 81; DB 1; Length 728;
Best Local Similarity 60.9%; Pred. No. 0.00084;
Matches 14; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```
Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||| |  ||| || ||   :||
Db    442 RNPDDDAHGPWCYTGNPLIPWDY 464
```

RESULT 16
TBBO
thrombin (EC 3.4.21.5) precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
C;Accession: S02537; A00915; A37552; I46045; S67518
R;Irwin, D.M.; Robertson, K.A.; MacGillivray, R.T.A.
J. Mol. Biol. 200, 31-45, 1988
A;Title: Structure and evolution of the bovine prothrombin gene.
A;Reference number: S02537; MUID:88245190
A;Accession: S02537
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-625 <IRW>
R;MacGillivray, R.T.A.; Davie, E.W.
Biochemistry 23, 1626-1634, 1984
A;Title: Characterization of bovine prothrombin mRNA and its translation product.
A;Reference number: A00915; MUID:84203525
A;Accession: A00915
A;Molecule type: mRNA
A;Residues: 1-230,'H',232-625 <MAC>
A;Note: 600-Asn was also found
R;Magnusson, S.; Sottrup-Jensen, L.; Petersen, T.E.; Claeys, H.
in Boerhaave Symposium on Prothrombin and Related Coagulation Factors, Hemker, H.C.,
A;Reference number: A37552
A;Accession: A37552
A;Molecule type: protein
A;Residues: 44-287,'N',289-352,'E',354,'Q',356-548,'ND',551-599,'N',601-625 <MAG>
A;Note: the evidence for 231-Ser is strong
A;Note: disulfide bonds and carbohydrate binding sites were determined
R;Park, C.H.; Tulinsky, A.
Biochemistry 25, 3977-3982, 1986
A;Title: Three-dimensional structure of the kringle sequence: structure of prothrombi
A;Reference number: A37553; MUID:86296631
A;Contents: annotation; residues 44-317, X-ray crystallography, 2.8 angstroms
R;Irwin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.
Biochemistry 24, 6854-6861, 1985
A;Title: Characterization of the bovine prothrombin gene.
A;Reference number: A37554; MUID:86077733
A;Contents: annotation; gene structure
R;MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980
A;Title: Cloning and analysis of a cDNA coding for bovine prothrombin.
A;Reference number: I46045; MUID:81054926
A;Accession: I46045
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 466-599,'N',601-625 <MA2>
A;Cross-references: EMBL:V00135; NID:g772; PIDN:CAA23451.1; PID:g808945
R;Pejler, G.; Karlstroem, A.R.; Berg, L.
Eur. J. Biochem. 227, 102-107, 1995
A;Title: Identification of the proteolytic thrombin fragments formed after cleavage w
A;Reference number: S67518; MUID:95154277
A;Accession: S67518
A;Status: preliminary
A;Molecule type: protein
A;Residues: 318-325;333-338,'X',340;367-374;481-484,'X',486-488;515-522 <PEJ>
C;Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fi
C;Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi
tivation peptide and cleaves the remaining part into light and heavy chains. The acti
.
C;Comment: Thrombin can cleave the amino-terminal activation peptide 1 from prothromb
C;Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carb
ent interaction with the negatively charged phospholipid membrane surface.
C;Comment: The prothrombin precursor is synthesized in the liver.
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; duplication; gl
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-625/Product: prothrombin #status experimental <MPT>
F;44-199/Domain: activation peptide 1 #status experimental <FR1>
F;109-187/Domain: kringle homology <KR1>
F;200-317/Domain: activation peptide 2 #status experimental <FR2>
F;214-292/Domain: kringle homology <KR2>

```
F;318-366/Product: thrombin light chain #status experimental <LCH>
F;367-625/Product: thrombin heavy chain #status experimental <HCH>
F;367-616/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,214-292,235-275,263-287,339-485,394-410,539-553,5
F;120,144,419/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;409,465,571/Active site: His, Asp, Ser #status experimental

  Query Match             58.0%;  Score 80;  DB 1;  Length 625;
  Best Local Similarity   70.6%;  Pred. No. 0.00099;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           ||||| : ||| |||:|
Db     159 RNPDGSITGPWCYTTSP 175


RESULT 17
I51283
hepatocyte growth factor precursor - clawed frog
N;Alternate names: hepapoietin A; scatter factor
C;Species: Xenopus sp. (clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I51283
R;Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A;Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A;Reference number: I51283; MUID:95267690
A;Accession: I51283
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-710 <NAK>
A;Cross-references: GB:S77422; NID:g998932; PIDN:AAB34354.1; PID:g998933
A;Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleoti
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F;42-477,478-709/Product: hepatocyte growth factor #status predicted <MAT>
F;42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;115-193/Domain: kringle homology <KR1>
F;198-275/Domain: kringle homology <KR2>
F;289-367/Domain: kringle homology <KR3>
F;375-453/Domain: kringle homology <KR4>
F;478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;478-700/Domain: trypsin homology <TRY>
F;52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre
F;470-588/Disulfide bonds: #status predicted

  Query Match             56.5%;  Score 78;  DB 1;  Length 710;
  Best Local Similarity   56.5%;  Pred. No. 0.0022;
  Matches   13;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || :|   :||
Db     425 RNPDNDAHGPWCYTDDPFVPWDY 447


RESULT 18
A40522
plasmin (EC 3.4.21.7) precursor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C;Accession: A40522
R;Kanalas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A;Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor s
A;Reference number: A40522; MUID:91250378
A;Accession: A40522
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-169 <KAN>
A;Cross-references: GB:M62832; NID:g206215; PIDN:AAA41884.1; PID:g554488
A;Note: the authors translated the codon TCT for residue 76 as Ala
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;34-112/Domain: kringle homology <KRG>
F;34-112,55-95,83-107/Disulfide bonds: #status predicted

  Query Match             55.8%;  Score 77;  DB 2;  Length 169;
  Best Local Similarity   52.2%;  Pred. No. 0.00066;
  Matches   12;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| :||:|   ::|
Db      84 RNPDNDQRGPWCFTTDPSVRWEY 106


RESULT 19
I51285
hepatocyte growth factor/scatter factor - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51285
R;Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gher
Development 121, 813-824, 1995
A;Title: A role for HGF/SF in neural induction and its expression in Hensen's node du
A;Reference number: I51285; MUID:95237013
A;Accession: I51285
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-411 <STR>
A;Cross-references: GB:S77480; NID:g998675; PID:g998676
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F;124-197/Domain: kringle homology <KRG>
F;202-279/Domain: kringle homology <KR2>
F;296-374/Domain: kringle homology <KR3>

  Query Match             54.3%;  Score 75;  DB 2;  Length 411;
  Best Local Similarity   50.0%;  Pred. No. 0.0034;
  Matches   11;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :||:|:  ::
Db     169 RNPRGEEGGPWCFTTSPQMRHE 190


RESULT 20
S33879
plasmin precursor - lamprey (fragments)
N;Alternate names: plasminogen
C;Species: Petromyzontidae gen. sp. (lamprey)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 07-Nov-1997
C;Accession: S33879
R;Affolter, M.; Schaller, J.; Rickli, E.E.
Protein Seq. Data Anal. 5, 207-211, 1993
A;Title: Isolation, characterization and partial amino acid sequence of lamprey plasm
A;Reference number: S33879
A;Accession: S33879
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15;16-34;35-44;45-59;60-76;77-111;111-138;139-158;159-178;179-216;217-2
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
F;81-146/Domain: kringle homology <KR3>

  Query Match             52.9%;  Score 73;  DB 2;  Length 336;
  Best Local Similarity   70.6%;  Pred. No. 0.0053;
  Matches   12;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           |||| |  ||| || ||
Db     189 RNPDQDPQGPWCYTXNP 205
```

```
RESULT 21
S10511
thrombin (EC 3.4.21.5) B chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 22-Jun-1999
C;Accession: S10511; A60576; B42696
R;Dihanich, M.; Monard, D.
Nucleic Acids Res. 18, 4251, 1990
A;Title: cDNA sequence of rat prothrombin.
A;Reference number: S10511; MUID:90332426
A;Accession: S10511
A;Molecule type: mRNA
A;Residues: 1-617 <DIH>
A;Cross-references: EMBL:X52835; NID:g56969; PIDN:CAA37017.1; PID:g56970
R;Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.
Endocrinology 126, 167-175, 1990
A;Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
A;Reference number: A60576; MUID:90091942
A;Accession: A60576
A;Molecule type: protein
A;Residues: 44-58 <HEN>
A;Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat ute
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A;Reference number: A42696; MUID:92212913
A;Accession: B42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 383-617,'E' <BAN>
A;Cross-references: GB:M81397
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-617/Product: prothrombin #status experimental <PMAT>
F;109-187/Domain: kringle homology <KR1>
F;215-292/Domain: kringle homology <KR2>
F;360-609/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,5
F;402,458,564/Active site: His, Asp, Ser #status predicted

  Query Match             52.9%;  Score 73;  DB 2;  Length 617;
  Best Local Similarity   64.7%;  Pred. No. 0.01;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           ||||    ||| |||:|
Db     159 RNPDSSTSGPWCYTTDP 175


RESULT 22
T18518
apolipoprotein(a) - western European hedgehog (fragment)
C;Species: Erinaceus europaeus (western European hedgehog)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C;Accession: T18518
R;Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A;Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprot
A;Reference number: I46259; MUID:96025778
A;Accession: T18518
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2869 <LAW>
A;Cross-references: EMBL:U33170; NID:g1046358; PID:g1046359; PIDN:AAC48522.1
A;Experimental source: liver
C;Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, con
ent apolipoprotein(a).

  Query Match             52.5%;  Score 72.5;  DB 2;  Length 2869;
  Best Local Similarity   56.5%;  Pred. No. 0.065;
  Matches   13;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||:|  || ||||    ::|
Db    2548 RNPDGEV-APWCYTTNSAMRWEY 2569


RESULT 23
A35827
thrombin (EC 3.4.21.5) B chain precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 22-Jun-1999
C;Accession: A35827; A42696; S12081
R;Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J
DNA Cell Biol. 9, 487-498, 1990
A;Title: Characterization of the cDNA coding for mouse prothrombin and localization o
A;Reference number: A35827; MUID:91025551
A;Accession: A35827
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-618 <DEG>
A;Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814
A;Experimental source: strain C57BL/6
A;Note: the data were obtained from females resulting from the cross of M. domesticus
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A;Reference number: A42696; MUID:92212913
A;Accession: A42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 384-618,'E' <BAN>
A;Cross-references: GB:M81394
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; h
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-618/Product: prothrombin B #status predicted <MAT>
F;109-187/Domain: kringle homology <KR1>
F;215-293/Domain: kringle homology <KR2>
F;361-610/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-54
F;403,459,565/Active site: His, Asp, Ser #status predicted

  Query Match             51.4%;  Score 71;  DB 2;  Length 618;
  Best Local Similarity   64.7%;  Pred. No. 0.02;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           ||||    ||| |||:|
Db     159 RNPDSSTTGPWCYTTDP 175


RESULT 24
TBHU
thrombin (EC 3.4.21.5) precursor [validated] - human
N;Alternate names: coagulation factor II
N;Contains: prothrombin
C;Species: Homo sapiens (man)
C;Date: 30-Nov-1980 #sequence_revision 22-Jul-1994 #text_change 08-Dec-2000
C;Accession: A29351; A00914; B00914; A37549; A37550; I51952
R;Degen, S.J.F.; Davie, E.W.
Biochemistry 26, 6165-6177, 1987
A;Title: Nucleotide sequence of the gene for human prothrombin.
A;Reference number: A29351; MUID:88077877
A;Accession: A29351
A;Molecule type: DNA
A;Residues: 1-622 <DEG>
A;Cross-references: GB:M17262; GB:M33691; NID:g558069; PIDN:AAC63054.1; PID:g339641
R;Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.
```

Biochemistry 22, 2087-2097, 1983
A;Title: Characterization of the complementary deoxyribonucleic acid and gene coding for
A;Reference number: A00914; MUID:83231469
A;Accession: A00914
A;Molecule type: mRNA
A;Residues: 8-163,'N',165-622 <DE2>
A;Cross-references: GB:V00595; GB:J00307; NID:g37128; PIDN:CAA23842.1; PID:g1335344
A;Accession: B00914
A;Molecule type: DNA
A;Residues: 188-311 <DE3>
R;Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.
Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977
A;Reference number: A37549; MUID:77193964
A;Accession: A37549
A;Molecule type: protein
A;Residues: 44-118,'N',120,'S',122-163,'I',165-175,'A',177-182,'T',184-193,'MV',196-308,
R;Butkowski, R.J.; Elion, J.; Downing, M.R.; Mann, K.G.
J. Biol. Chem. 252, 4942-4957, 1977
A;Title: Primary structure of human prethrombin 2 and alpha-thrombin.
A;Reference number: A37550; MUID:77207112
A;Accession: A37550
A;Molecule type: protein
A;Residues: 315-334,'N',336-348,'N',350-368,'N',370-397,'N',399-413,'N',415-484,'N',486-
R;Rabiet, M.J.; Blashill, A.; Furie, B.; Furie, B.C.
J. Biol. Chem. 261, 13210-13215, 1986
A;Reference number: A37551; MUID:87008532
A;Contents: annotation; activation cleavages
R;MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.
Ann. N. Y. Acad. Sci. 485, 73-79, 1986
A;Title: Recombinant genetic approaches to functional mapping of thrombin.
A;Reference number: I51952; MUID:87182874
A;Accession: I51952
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2,'RI',5-100 <RES>
A;Cross-references: GB:M33031; NID:g190723; PIDN:AAA60220.1; PID:g190724
C;Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibri
C;Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds
) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chai
ter 314-Arg, are released in natural blood clotting.
C;Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.
C;Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxy
ent interaction with the negatively charged phospholipid membrane surface.
C;Comment: The prothrombin precursor is synthesized in the liver.
C;Genetics:
A;Gene: GDB:F2
A;Cross-references: GDB:119894; OMIM:176930
A;Map position: 11p11-11q12
A;Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; dupli
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-87/Domain: Gla domain homology <GLA>
F;44-622/Product: prothrombin #status experimental <MAT>
F;44-327/Domain: activation peptide #status experimental <APT>
F;108-186/Domain: kringle homology <KR1>
F;213-291/Domain: kringle homology <KR2>
F;328-363/Product: thrombin light chain #status experimental <LCH>
F;364-622/Product: thrombin heavy chain #status experimental <HCH>
F;364-613/Domain: trypsin homology <TRY>
F;49,50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds: #status
F;121,143/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;336-482,536-550,564-594/Disulfide bonds: #status predicted
F;391-407/Disulfide bonds: #status experimental
F;406,462/Active site: His, Asp #status predicted
F;416/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;568/Active site: Ser #status experimental

Query Match 51.4%; Score 71; DB 1; Length 622;
Best Local Similarity 64.7%; Pred. No. 0.021;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RNPDGDVGGPWAYTTNP 17
|||| ||| |||:|
Db 158 RNPDSSTTGPWCYTTDP 174

RESULT 25
C61545
plasmin (EC 3.4.21.7) precursor - goat (fragments)
N;Alternate names: plasminogen
C;Species: Capra aegagrus hircus (domestic goat)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C;Accession: C61545
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: C61545
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-123 <SCH>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: hydrolase; serine proteinase
F;41-118/Domain: kringle homology <KR4>

Query Match 49.6%; Score 68.5; DB 2; Length 123;
Best Local Similarity 52.2%; Pred. No. 0.0082;
Matches 12; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
|||| | || |||:|| :::
Db 91 RNPDAD-KSPWCYTTDPRVRWEF 112

RESULT 26
A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N;Alternate names: plasminogen
C;Species: Gallus gallus (chicken)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C;Accession: A60140
R;Gyenes, M.; Patthy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A;Title: The kringle 4 domain of chicken plasminogen.
A;Reference number: A60140; MUID:86077796
A;Accession: A60140
A;Molecule type: protein
A;Residues: 1-89 <GYE>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;6-83/Domain: kringle homology <KRG>
F;6-83,27-66,55-78/Disulfide bonds: #status predicted
F;39/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 48.2%; Score 66.5; DB 2; Length 89;
Best Local Similarity 52.2%; Pred. No. 0.011;
Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 1 RNPDGDVGGPWAYTTNPRKLYDY 23
|||| | || |||:| ::|
Db 56 RNPDAD-RSPWCYTTDPSVRWEY 77

RESULT 27
A32869
apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C;Accession: A32869; A30848
R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis
A;Reference number: A32869; MUID:89174660

```
A;Accession: A32869
A;Molecule type: mRNA
A;Residues: 1-1420 <TOM>
A;Cross-references: GB:J04635; NID:g342072; PIDN:AAA36833.1; PID:g342073
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;50-127/Domain: kringle homology <KR1>
F;164-241/Domain: kringle homology <KR2>
F;278-355/Domain: kringle homology <KR3>
F;392-469/Domain: kringle homology <KR4>
F;506-583/Domain: kringle homology <KR5>
F;620-697/Domain: kringle homology <KR6>
F;726-803/Domain: kringle homology <KR7>
F;840-917/Domain: kringle homology <KR8>
F;954-1031/Domain: kringle homology <KR9>
F;1068-1145/Domain: kringle homology <KR10>
F;1191-1413/Domain: trypsin homology <TRY>

  Query Match             47.5%;  Score 65.5;  DB 2;  Length 1420;
  Best Local Similarity   47.8%;  Pred. No. 0.32;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| ::  || || :||  ::|
Db     776 RNPDAEI-RPWCYTMDPRVRWEY 797


RESULT 28
E61545
plasmin (EC 3.4.21.7) precursor - dog (fragments)
N;Alternate names: plasminogen
C;Species: Canis lupus familiaris (dog)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C;Accession: E61545
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: E61545
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-120 <SCH>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: hydrolase; serine proteinase
F;37-114/Domain: kringle homology <KR4>

  Query Match             45.3%;  Score 62.5;  DB 2;  Length 120;
  Best Local Similarity   47.8%;  Pred. No. 0.06;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |   || |||:|   :::
Db      87 RNPDAD-KSPWCYTTDPSVRWEF 108


RESULT 29
E70711
hypothetical protein Rv1490 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: E70711
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: E70711
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-435 <COL>
A;Cross-references: GB:Z79701; GB:AL123456; NID:g3261635; PIDN:CAB02040.1; PID:g1524237
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1490
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1490

  Query Match             43.5%;  Score 60;  DB 2;  Length 435;
  Best Local Similarity   47.6%;  Pred. No. 0.57;
  Matches   10;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       3 PDGDVGGPWAYTTNPRKLYDY 23
           | | |  |||  | |::: ||
Db     259 PHGMVENPWAVATTPQRVLDY 279


RESULT 30
B45082
neurotrophic receptor ror2 precursor - human
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
C;Accession: B45082
R;Masiakowski, P.; Carroll, R.D.
J. Biol. Chem. 267, 26181-26190, 1992
A;Title: A novel family of cell surface receptors with tyrosine kinase-like domain.
A;Reference number: A45082; MUID:93100347
A;Accession: B45082
A;Molecule type: mRNA
A;Residues: 1-943 <MAS>
A;Cross-references: GB:M97639; NID:g337466; PIDN:AAA60276.1; PID:g337467
A;Note: sequence extracted from NCBI backbone (NCBIP:120918)
C;Genetics:
A;Gene: GDB:NTRKR2
A;Cross-references: GDB:136454
A;Map position: 6p21-6p21
C;Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology;
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; ty
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-937/Product: neurotrophic receptor ror2 #status predicted <MAT>
F;76-137/Domain: immunoglobulin homology <IMM>
F;316-394/Domain: kringle homology <KRG>
F;412-428/Domain: transmembrane #status predicted <TMN>
F;471-753/Domain: protein kinase homology <KIN>
F;479-487/Region: protein kinase ATP-binding motif
F;70,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             42.8%;  Score 59;  DB 2;  Length 943;
  Best Local Similarity   56.2%;  Pred. No. 1.8;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTN 16
           ||| | : ||| :| |
Db     366 RNPGGQMEGPWCFTQN 381


RESULT 31
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
R;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. S
A;Reference number: A35029; MUID:90130448
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activ
A;Reference number: A31597; MUID:89170114
```

```
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379,'K',381-559 <NYT>
A;Cross-references: GB:M23697; NID:g530159; PIDN:AAA41812.1; PID:g530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

  Query Match             40.6%;  Score 56;  DB 1;  Length 559;
  Best Local Similarity   50.0%;  Pred. No. 2.9;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    2;  Gaps    2;

Qy      1 RNPDGDVGGPWAYTTNPRKL-YDY 23
          ||||||   || :    ||| ::|
Db    266 RNPDGD-AKPWCHVMKDRKLTWEY 288


RESULT 32
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29941; S48205; S48207; S48206
R;Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR
A;Reference number: A29941; MUID:88087303
A;Accession: A29941
A;Molecule type: mRNA
A;Residues: 1-559 <RIC>
A;Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076
A;Accession: S48205
A;Molecule type: protein
A;Residues: 33-37,'X',39-40 <LIJ>
A;Accession: S48207
A;Molecule type: protein
A;Residues: 309-316 <LI2>
A;Accession: S48206
A;Molecule type: protein
A;Residues: 33-37,'X',39-40 <LIW>
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

  Query Match             40.6%;  Score 56;  DB 1;  Length 559;
  Best Local Similarity   50.0%;  Pred. No. 2.9;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    2;  Gaps    2;

Qy      1 RNPDGDVGGPWAYTTNPRKL-YDY 23
          ||||||   || :    ||| ::|
Db    266 RNPDGD-ARPWCHVMKDRKLTWEY 288


RESULT 33
T37219
probable lipoprotein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37219
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21615
A;Accession: T37219
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-153 <OLI>
A;Cross-references: EMBL:AL031514; PIDN:CAA20597.1; GSPDB:GN00070; SCOEDB:SC2H4.03
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC2H4.03

  Query Match             38.8%;  Score 53.5;  DB 2;  Length 153;
  Best Local Similarity   44.0%;  Pred. No. 1.6;
  Matches   11;  Conservative    1;  Mismatches    4;  Indels    9;  Gaps    1;

Qy      3 PDGDVGGPWA---------YTTNPR 18
          | || || ||          | |:||
Db     77 PGGDTGGDWAFFADDGDGLYMTSPR 101


RESULT 34
T25094
hypothetical protein T22A3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T25094
R;McMurray, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19980
A;Accession: T25094
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-323 <WIL>
A;Cross-references: EMBL:Z81125; PIDN:CAB03383.1; GSPDB:GN00019; CESP:T22A3.6
A;Experimental source: clone T22A3
C;Genetics:
A;Gene: CESP:T22A3.6
A;Map position: 1
A;Introns: 104/2; 136/2; 226/3; 242/2
C;Superfamily: Caenorhabditis elegans hypothetical protein T22A3.6

  Query Match             37.7%;  Score 52;  DB 2;  Length 323;
  Best Local Similarity   56.2%;  Pred. No. 6.1;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTN 16
          |||| :  ||| |  |
Db     14 RNPDKNPLGPWCYVGN 29


RESULT 35
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N;Alternate names: t-PA; tissue plasminogen activator
C;Species: Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
```

C;Accession: A94004; A23529; JT0562; A93293; S02125; A91343; A93951; A91322; A54645; I60
R;Ny, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation
A;Reference number: A94004; MUID:84298137
A;Accession: A94004
A;Molecule type: DNA
A;Residues: 1-562 <NYT>
A;Cross-references: GB:L00141
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translati
R;Friezner Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A;Title: The human tissue plasminogen activator gene.
A;Reference number: A23529; MUID:86196143
A;Accession: A23529
A;Molecule type: DNA
A;Residues: 1-562 <DEG>
A;Cross-references: GB:K03021; NID:g339817; PIDN:AAA98809.1; PID:g339818
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A;Title: Purification and characterization of tissue plasminogen activator secreted by h
A;Reference number: JT0562; MUID:91291340
A;Accession: JT0562
A;Molecule type: mRNA
A;Residues: 31-562 <ITA>
A;Cross-references: DDBJ:D01096; NID:g220128; PIDN:BAA00881.1; PID:g441174
A;Experimental source: embryonic lung fibroblast IMR-90 cells
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A;Reference number: A93293; MUID:83115262
A;Accession: A93293
A;Molecule type: mRNA
A;Residues: 1-562 <PEN>
A;Cross-references: GB:L00141
A;Experimental source: melanoma cells
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A;Reference number: S02125; MUID:88262579
A;Accession: S02125
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-562 <SAS>
A;Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244
A;Experimental source: fetal lung cells
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
FEBS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; MUID:85285620
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38,'G',86-433,'E',435-562 <KAG>
A;Experimental source: Detroit 562 cells; ATCC 138
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
A;Reference number: A93951; MUID:83169656
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <EDL>
A;Experimental source: melanoma cells
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A;Reference number: A90488; MUID:85000468
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A;Reference number: A91322; MUID:84158956
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45;311-320 <POH>
A;Experimental source: uterus
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; MUID:87033611
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; En
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
A;Reference number: A37568; MUID:87161761
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-typ
A;Reference number: A60902; MUID:89044681
A;Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its e
A;Reference number: A54645; MUID:86284200
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A;Note: parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A;Reference number: I60110; MUID:88054470
A;Accession: I60110
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-562 <RES>
A;Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activato
A;Reference number: I55232; MUID:85289338
A;Accession: I55232
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a sing
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repea
C;Genetics:
A;Gene: GDB:PLAT
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3;
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-562/Product: t-plasminogen activator #status experimental <MAT>
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F;41-78/Domain: fibronectin type I repeat homology <1F1>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-296/Domain: kringle homology <KR2>
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F;311-556/Domain: trypsin homology <TRY>
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,29
F;152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F;357,406/Active site: His, Asp #status predicted
F;513/Active site: Ser #status experimental

```
  Query Match             37.7%;  Score 52;  DB 1;  Length 562;
  Best Local Similarity   45.8%;  Pred. No. 11;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    2;  Gaps    2;

Qy      1 RNPDGDVGGPWAYTTNPRKL-YDY 23
          ||||||   || :    |:| ::|
Db    268 RNPDGD-AKPWCHVLKNRRLTWEY 290


RESULT 36
T10851
y4hQ protein - Rhizobium sp. (strain NGR234) plasmid pNGR234a
C;Species: Rhizobium sp.
A;Variety: strain NGR234
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: T10851
R;Freiberg, C.; Fellay, R.; Bairoch, A.; Broughton, W.J.; Rosenthal, A.; Perret, X.
Nature 387, 394-401, 1997
A;Title: Molecular basis of symbiosis between Rhizobium and legumes.
A;Reference number: Z14734; MUID:97305956
A;Accession: T10851
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-194 <FRE>
A;Cross-references: EMBL:AE000077; NID:g2182428; PID:g2182436
C;Genetics:
A;Gene: y4hQ
A;Genome: plasmid pNGR234a

  Query Match             37.0%;  Score 51;  DB 2;  Length 194;
  Best Local Similarity   72.7%;  Pred. No. 4.9;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 PDGDVGGPWAY 13
          |  ||||||  |
Db    128 PPEDVGGPWGY 138


RESULT 37
S36786
carboxylesterase (EC 3.1.1.1) E4 - green peach aphid
C;Species: Myzus persicae (green peach aphid)
C;Date: 09-Jun-1994 #sequence_revision 01-Dec-1995 #text_change 18-Jun-1999
C;Accession: S36786
R;Field, L.M.; Williamson, M.S.; Moores, G.D.; Devonshire, A.L.
Biochem. J. 294, 569-574, 1993
A;Title: Cloning and analysis of the esterase genes conferring insecticide resistance in
A;Reference number: S36786; MUID:93384534
A;Accession: S36786
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-552 <FIE>
A;Cross-references: EMBL:X74554; NID:g397510; PIDN:CAA52648.1; PID:g397511
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase
F;52-552/Domain: cholinesterase homology <CHE>

  Query Match             37.0%;  Score 51;  DB 2;  Length 552;
  Best Local Similarity   42.1%;  Pred. No. 15;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 GDVGGPWAYTTNPRKLYDY 23
          |:: | : || | ||:| :
Db     34 GEIAGGFEYTYNGRKIYSF 52


RESULT 38
S36787
carboxylesterase (EC 3.1.1.1) FE4 precursor - green peach aphid
C;Species: Myzus persicae (green peach aphid)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S36787
R;Field, L.M.; Williamson, M.S.; Moores, G.D.; Devonshire, A.L.
Biochem. J. 294, 569-574, 1993
A;Title: Cloning and analysis of the esterase genes conferring insecticide resistance
A;Reference number: S36786; MUID:93384534
A;Accession: S36787
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-564 <FIE>
A;Cross-references: EMBL:X74555; NID:g397512; PIDN:CAA52649.1; PID:g397513
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase
F;52-552/Domain: cholinesterase homology <CHE>

  Query Match             37.0%;  Score 51;  DB 1;  Length 564;
  Best Local Similarity   42.1%;  Pred. No. 16;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 GDVGGPWAYTTNPRKLYDY 23
          |:: | : || | ||:| :
Db     34 GEIAGGFEYTYNGRKIYSF 52


RESULT 39
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0600
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.;
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat
A;Reference number: JS0597; MUID:92039036
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:g166078; PIDN:AAA31595.1; PID:g166079
A;Note: the authors translated the codon ATC for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>
F;45-126/Domain: kringle homology <KRG>
F;143-388/Domain: trypsin homology <TRY>
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bond
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             36.6%;  Score 50.5;  DB 2;  Length 394;
  Best Local Similarity   52.6%;  Pred. No. 13;
  Matches   10;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAYTTNPRK 19
          |||||    || |    ||
Db     98 RNPDG-ASKPWCYVIKARK 115


RESULT 40
E83194
polyamine transport protein PotB PA3608 [imported] - Pseudomonas aeruginosa (strain P
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83194
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: E83194
A;Status: preliminary
A;Molecule type: DNA
```

A;Residues: 1-297 <STO>
A;Cross-references: GB:AE004781; GB:AE004091; NID:g9949761; PIDN:AAG06996.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: potB; PA3608
C;Superfamily: spermidine/putrescine transport system permease protein potH

Query Match 35.5%; Score 49; DB 2; Length 297;
Best Local Similarity 45.0%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 5 GDVGG--PWAYTTNPRKLYD 22
|| || || |: : : :||
Db 38 GDYGGLAPWMYSEDGQTVYD 57

RESULT 41
G70520
probable csp protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: G70520
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: G70520
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-539 <COL>
A;Cross-references: GB:Z97188; GB:AL123456; NID:g3261805; PIDN:CAB10019.1; PID:e1300066;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: csp

Query Match 35.5%; Score 49; DB 2; Length 539;
Best Local Similarity 52.4%; Pred. No. 29;
Matches 11; Conservative 1; Mismatches 5; Indels 4; Gaps 2;

Qy 1 RNPDGDVGGPW---AYTTNPR 18
: |||| ||| | | ||
Db 72 KRPDGD-WGPWYQTEYETEPR 91

RESULT 42
KFHU12
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N;Alternate names: Hageman factor (activated)
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037
R;Cool, D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon ge
A;Reference number: A29411; MUID:88007593
A;Accession: A29411
A;Molecule type: DNA
A;Residues: 1-615 <COO>
A;Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AAB59490.1; PID:g180357
R;Tripodi, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A;Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A;Reference number: A26814; MUID:86176794
A;Accession: A26814
A;Molecule type: mRNA
A;Residues: 4-615 <TRI>
A;Cross-references: GB:M31315; NID:g182291; PIDN:AAA70225.1; PID:g182292
R;Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.T
J. Biol. Chem. 260, 13666-13676, 1985
A;Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
A;Reference number: A00930; MUID:86033830
A;Accession: A00930
A;Molecule type: mRNA
A;Residues: 14-332,'S',334-615 <CO2>
A;Cross-references: GB:M11723; NID:g180358; PIDN:AAA51986.1; PID:g180359
R;Que, B.G.; Davie, E.W.
Biochemistry 25, 1525-1528, 1986
A;Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A;Reference number: A25191; MUID:86216049
A;Accession: A25191
A;Molecule type: mRNA
A;Residues: 146-378,'G',380-615 <QUE>
A;Cross-references: GB:M13147; NID:g180360; PIDN:AAA70224.1; PID:g180361
R;McMullen, B.A.; Fujikawa, K.
J. Biol. Chem. 260, 5328-5341, 1985
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated
A;Reference number: A22248; MUID:85182674
A;Accession: A22248
A;Molecule type: protein
A;Residues: 20-379 <MCM>
R;Fujikawa, K.; McMullen, B.A.
J. Biol. Chem. 258, 10924-10933, 1983
A;Title: Amino acid sequence of human beta-factor XIIa.
A;Reference number: A21037; MUID:83291041
A;Accession: A21037
A;Molecule type: protein
A;Residues: 354-362;373-615 <FUJ>
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of fa
A;Reference number: A44606; MUID:92184750
A;Contents: annotation; carbohydrate binding site
C;Genetics:
A;Gene: GDB:F12
A;Cross-references: GDB:119892; OMIM:234000
A;Map position: 5q34-5qter
A;Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 5
C;Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to anion
C;Function:
A;Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasm
ikrein
A;Pathway: blood coagulation; fibrinolysis
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolo
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A
F;47-88/Domain: fibronectin type II repeat homology <FB2>
F;98-130/Domain: EGF homology <EG1>
F;135-170/Domain: fibronectin type I repeat homology <1F1>
F;178-209/Domain: EGF homology <EG2>
F;217-295/Domain: kringle homology <KRG>
F;298-356/Region: proline-rich
F;354-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B
F;373-609/Domain: trypsin homology <TRY>
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;299,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;412,461,563/Active site: His, Asp, Ser #status predicted

Query Match 35.5%; Score 49; DB 1; Length 615;
Best Local Similarity 41.7%; Pred. No. 34;
Matches 10; Conservative 5; Mismatches 7; Indels 2; Gaps 2;

Qy 1 RNPDGDVGGPWAYTTNPRKL-YDY 23
|||| |: || : | :| ::|
Db 267 RNPDNDI-RPWCFVLNRDRLSWEY 289

RESULT 43
A47299
ror-related receptor RTK - Pacific electric ray

```
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Torpedo californica (Pacific electric ray)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-May-2000
C;Accession: A47299
R;Jennings, C.G.; Dyer, S.M.; Burden, S.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 2895-2899, 1993
A;Title: Muscle-specific trk-related receptor with a kringle domain defines a distinct c
A;Reference number: A47299; MUID:93219391
A;Accession: A47299
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-946 <JEN>
A;Cross-references: GB:L11311; NID:g290857; PIDN:AAA49285.1; PID:g290858
A;Experimental source: electric organ
A;Note: sequence extracted from NCBI backbone (NCBIN:128724, NCBIP:128726)
C;Superfamily: Torpedo ror-related receptor; immunoglobulin homology; kringle homology;
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyros
F;44-103/Domain: immunoglobulin homology <IM1>
F;137-195/Domain: immunoglobulin homology <IM2>
F;229-287/Domain: immunoglobulin homology <IM3>
F;464-542/Domain: kringle homology <KR3>
F;572-588/Domain: transmembrane #status predicted <TMN>
F;650-940/Domain: protein kinase homology <KIN>
F;658-666/Region: protein kinase ATP-binding motif
F;225,340,477,544/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             35.5%;  Score 49;  DB 1;  Length 946;
  Best Local Similarity   57.1%;  Pred. No. 54;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYT 14
          ||| |:   || ||
Db    514 RNPGGESERPWCYT 527


RESULT 44
T49228
plasma membrane H+-ATPase-like - Arabidopsis thaliana
N;Alternate names: protein F27H5.120
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C;Accession: T49228
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25018
A;Accession: T49228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-961 <RIE>
A;Cross-references: EMBL:AL163852; GSPDB:GN00061; ATSP:F27H5.120
A;Experimental source: cultivar Columbia; BAC clone F27H5
C;Genetics:
A;Gene: ATSP:F27H5.120
A;Map position: 3
A;Introns: 17/3; 57/3; 90/3; 136/3; 183/1; 241/3; 310/3; 350/3; 390/3; 431/3; 466/3; 515
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F;484-653/Domain: ATPase nucleotide-binding domain homology <ATN>

  Query Match             35.1%;  Score 48.5;  DB 2;  Length 961;
  Best Local Similarity   64.3%;  Pred. No. 65;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      3 PDGDV---GGPWAY 13
          |||||   |||| :
Db    469 PDGDVKGEGGPWDF 482


RESULT 45
B64318
hypothetical protein MJ0145 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 11-Jan-2000
C;Accession: B64318
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999
A;Accession: B64318
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-268 <BUL>
A;Cross-references: GB:U67471; GB:L77117; NID:g1590901; PID:g1498914; TIGR:MJ0145
C;Genetics:
A;Map position: REV142991-142185
C;Superfamily: Methanobacterium conserved hypothetical protein MTH1017

  Query Match             34.8%;  Score 48;  DB 2;  Length 268;
  Best Local Similarity   57.1%;  Pred. No. 19;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      9 GPWAYTTNPRKLYD 22
          |||  | |||:  |
Db     13 GPWTVTPNPRRESD 26


RESULT 46
T32768
hypothetical protein F33D11.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C;Accession: T32768
R;Sammons, L.; Wohldmann, P.; Mullen, G.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid F33D11.
A;Reference number: Z21222
A;Accession: T32768
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-347 <SAM>
A;Cross-references: EMBL:AF039720; PIDN:AAB96700.1; GSPDB:GN00019; CESP:F33D11.7
A;Experimental source: strain Bristol N2; clone F33D11
C;Genetics:
A;Gene: CESP:F33D11.7
A;Map position: 1
A;Introns: 34/3; 72/3; 104/2; 127/3; 159/3; 187/3; 260/3; 295/3; 324/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

  Query Match             34.8%;  Score 48;  DB 2;  Length 347;
  Best Local Similarity   42.9%;  Pred. No. 25;
  Matches    6;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     10 PWAYTTNPRKLYDY 23
          |||:   |::::||
Db    255 PWAHLKKPKEVHDY 268


RESULT 47
T32881
hypothetical protein C34B2.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T32881
R;Graves, T.; Sutterer, C.; Hawkins, M.; Wilson, R.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid C34B2.
A;Reference number: Z21241
A;Accession: T32881
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <GRA>
A;Cross-references: EMBL:AF043693; PIDN:AAB97535.1; GSPDB:GN00019; CESP:C34B2.3
A;Experimental source: strain Bristol N2; clone C34B2
C;Genetics:
```

```
A;Gene: CESP:C34B2.3
A;Map position: 1
A;Introns: 40/3; 79/3; 111/2; 134/3; 166/3; 194/3; 267/3; 302/3; 331/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

  Query Match             34.8%;  Score 48;  DB 2;  Length 357;
  Best Local Similarity   42.9%;  Pred. No. 26;
  Matches    6;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 PWAYTTNPRKLYDY 23
           |||:   |::::||
Db     262 PWAHLKKPKEVHDY 275


RESULT 48
AC3603
2-oxoisovalerate dehydrogenase alpha chain (EC 1.2.4.4) [imported] - Brucella melitensis
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AC3603
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3603
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-410 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53990.1; PID:g17984939; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0748
A;Map position: II
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin
C;Keywords: oxidoreductase

  Query Match             34.8%;  Score 48;  DB 2;  Length 410;
  Best Local Similarity   42.9%;  Pred. No. 30;
  Matches    9;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       2 NPDGDVGGPWAYTTNPRKLYD 22
           | :|:  |||| |    :| |
Db      59 NREGEAVGPWAGTLTDEELKD 79


RESULT 49
F70662
probable plcC protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: F70662
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: F70662
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-508 <COL>
A;Cross-references: GB:Z83860; GB:AL123456; NID:g3261681; PIDN:CAB06146.1; PID:g1781256
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: plcC

  Query Match             34.8%;  Score 48;  DB 2;  Length 508;
  Best Local Similarity   88.9%;  Pred. No. 38;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 NPDGDVGGP 10
           ||||| |||
Db     196 NPDGDQGGP 204


RESULT 50
H84099
cell wall-binding protein BH3600 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84099
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <STO>
A;Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BAB07319.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3600

  Query Match             34.1%;  Score 47;  DB 2;  Length 461;
  Best Local Similarity   40.0%;  Pred. No. 48;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    4;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNP----RKLY 21
           |   ||:  |: | |:|    |||:
Db     328 RPATGDISSPFGYRTHPVTGQRKLH 352


Search completed: November  8, 2002, 09:32:03
Job time : 17 secs
```

(9)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:31:21 ; Search time 26 Seconds
(without alignments)
153.034 Million cell updates/sec

Title: US-09-657-431-9
Perfect score: 138
Sequence: 1 RNPDGDVGGPWAYTTNPRKLYDY 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	97.1	334	6	O46507	O46507 papio hamad
2	134	97.1	810	4	Q15146	Q15146 homo sapien
3	128	92.8	812	11	Q91WJ5	Q91wj5 mus musculu
4	122	88.4	812	11	Q9R0W3	Q9r0w3 rattus norv
5	115	83.3	806	6	O18783	O18783 macropus eu
6	94	68.1	716	11	P70521	P70521 rattus norv
7	90	65.2	709	13	Q90ZN6	Q90zn6 brachydanio
8	90	65.2	716	11	Q91XG8	Q91xg8 mus musculu
9	88	63.8	567	4	Q13208	Q13208 homo sapien
10	86	62.3	313	13	Q9PU78	Q9pu78 crocodylus
11	86	62.3	704	13	Q90865	Q90865 gallus gall
12	84	60.9	648	4	Q9H1V4	Q9h1v4 homo sapien
13	81	58.7	728	6	Q9BH09	Q9bh09 felis silve
14	79	57.2	75	6	Q9BGN9	Q9bgn9 bos taurus
15	79	57.2	109	6	Q9N1B8	Q9n1b8 ovis aries
16	79	57.2	208	4	Q9BYM0	Q9bym0 homo sapien
17	79	57.2	210	4	Q13494	Q13494 homo sapien
18	79	57.2	211	11	O55027	O55027 mus musculu
19	79	57.2	290	4	Q02935	Q02935 homo sapien
20	79	57.2	296	4	Q14519	Q14519 homo sapien
21	79	57.2	716	13	Q91691	Q91691 xenopus lae
22	79	57.2	717	13	P70006	P70006 xenopus lae
23	78	56.5	385	5	Q25101	Q25101 herdmania m
24	78	56.5	710	13	Q91402	Q91402 xenopus. he
25	76	55.1	726	13	Q90978	Q90978 gallus gall
26	72.5	52.5	2869	6	Q28398	Q28398 erinaceus e
27	68	49.3	215	13	O42341	O42341 gallus gall
28	66.5	48.2	132	4	Q16609	Q16609 homo sapien
29	66	47.8	385	13	Q90WS2	Q90ws2 elaphe sp.
30	64.5	46.7	454	6	O46506	O46506 papio hamad
31	64	46.4	607	13	Q91001	Q91001 gallus gall
32	64	46.4	608	13	Q9PTW7	Q9ptw7 struthio ca
33	63.5	46.0	113	4	Q9UIR5	Q9uir5 homo sapien
34	63	45.7	378	13	Q90WP0	Q90wp0 trachemys s
35	62.5	45.3	105	4	Q9UIR8	Q9uir8 homo sapien
36	60	43.5	111	6	O77688	O77688 oryctolagus
37	60	43.5	286	4	Q13209	Q13209 homo sapien
38	59.5	43.1	60	4	Q9UKJ7	Q9ukj7 homo sapien
39	59.5	43.1	113	4	Q9UIR7	Q9uir7 homo sapien
40	59	42.8	420	13	Q90504	Q90504 eptatretus
41	56	40.6	263	4	O00318	O00318 homo sapien
42	56	40.6	263	4	Q96FE7	Q96fe7 homo sapien
43	56	40.6	559	11	Q91VP2	Q91vp2 mus musculu
44	55.5	40.2	113	4	Q9UIR6	Q9uir6 homo sapien
45	55	39.9	382	13	Q90WT4	Q90wt4 crocodylus
46	54.5	39.5	117	4	Q9UGS5	Q9ugs5 homo sapien
47	54.5	39.5	452	13	Q90Y90	Q90y90 xenopus lae
48	54.5	39.5	473	4	Q9BY70	Q9by70 homo sapien
49	54.5	39.5	473	11	Q99N43	Q99n43 mus musculu
50	54.5	39.5	473	11	Q924S4	Q924s4 rattus norv

ALIGNMENTS

RESULT 1
O46507
ID O46507 PRELIMINARY; PRT; 334 AA.
AC O46507;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PLASMINOGEN (FRAGMENT).
GN BABPEPSG.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF029692; AAB97887.1; -.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.

```
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  36791 MW;  C7DC06E03B965286 CRC64;

  Query Match             97.1%;  Score 134;  DB 6;  Length 334;
  Best Local Similarity   95.7%;  Pred. No. 1.3e-11;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db      56 RNPDGDVGGPWCYTTNPRKLYDY 78


RESULT 2
Q15146
ID   Q15146      PRELIMINARY;      PRT;   810 AA.
AC   Q15146;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA   Mitchell D., Robinson J.H.;
RT   "Expression of recombinant human plasminogen and aglycoplasminogen in
RT   HeLa cells.";
RL   Fibrinolysis 0:0-0(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; M74220; AAA36451.1; -.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    810       PLASMINOGEN.
SQ   SEQUENCE   810 AA;  90555 MW;  B05C7D4B0D020B3C CRC64;

  Query Match             97.1%;  Score 134;  DB 4;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 3.5e-11;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| ||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 3
Q91WJ5
ID   Q91WJ5      PRELIMINARY;      PRT;   812 AA.
AC   Q91WJ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014773; AAH14773.1; -.
SQ   SEQUENCE   812 AA;  90781 MW;  24173260E6A2FFD2 CRC64;

  Query Match             92.8%;  Score 128;  DB 11;  Length 812;
  Best Local Similarity   91.3%;  Pred. No. 2.6e-10;
  Matches   21;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| ||||||||||
Db     532 RNPDGDVNGPWCYTTNPRKLYDY 554


RESULT 4
Q9R0W3
ID   Q9R0W3      PRELIMINARY;      PRT;   812 AA.
AC   Q9R0W3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN PROTEIN PRECURSOR (EC 3.4.21.7).
GN   PLASMINOGEN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Bangert K., Johnsen A.H., Thorsen S.;
RT   "Rat plasminogen: cDNA and gene structure.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=91250378; PubMed=1645711;
RA   Kanalas J.J., Makker S.P.;
RT   "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT   receptor site for plasminogen.";
RL   J. Biol. Chem. 266:10825-10829(1991).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AJ242649; CAB46014.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
```

```
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    812       PLASMINOGEN.
SQ   SEQUENCE   812 AA;  90535 MW;  8C703C51410EBC9E CRC64;

  Query Match             88.4%;  Score 122;  DB 11;  Length 812;
  Best Local Similarity   87.0%;  Pred. No. 2e-09;
  Matches   20;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||| ||| || ||||||||
Db    532 RNPDGDVNGPWCYTMNPRKLYDY 554


RESULT 5
O18783
ID   O18783      PRELIMINARY;      PRT;   806 AA.
AC   O18783;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMINOGEN.
OS   Macropus eugenii (Tammar wallaby).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=98004511; PubMed=9342350;
RA   Lawn R.M., Schwartz K., Patthy L.;
RT   "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF012297; AAB65760.1; -.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   806 AA;  90981 MW;  95FAA86DC20064D5 CRC64;

  Query Match             83.3%;  Score 115;  DB 6;  Length 806;
  Best Local Similarity   78.3%;  Pred. No. 2.1e-08;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||  ||| |||||:||:||
Db    527 RNPDGDPNGPWCYTTNPKKLFDY 549


RESULT 6
P70521
ID   P70521      PRELIMINARY;      PRT;   716 AA.
AC   P70521;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MACROPHAGE STIMULATING PROTEIN PRECURSOR.
GN   MSP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=97011126; PubMed=8858136;
RA   Ohshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA   Takasu N., Suda T.;
RT   "Molecular cloning of Rat Macrophage-stimulating protein and its
RT   involvement in the Male Reproductive System.";
RL   Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; X95096; CAA64473.1; -.
DR   HSSP; P00747; 1KRN.
DR   MEROPS; S01.975; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     31       POTENTIAL.
SQ   SEQUENCE   716 AA;  80733 MW;  06B7DF3EF56D921F CRC64;

  Query Match             68.1%;  Score 94;  DB 11;  Length 716;
  Best Local Similarity   65.2%;  Pred. No. 2.2e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||  ||| || :|  |:||
Db    429 RNPDGDSHGPWCYTLDPETLFDY 451


RESULT 7
Q90ZN6
ID   Q90ZN6      PRELIMINARY;      PRT;   709 AA.
AC   Q90ZN6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE 1.
GN   HGFL1.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bassett D.I., Wilson S.W.;
RT   "Early expression of zebrafish Hepatocyte Growth Factor-Like 1
RT   suggests a conserved role in vertebrate neural induction.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF370035; AAK54207.1; -.
SQ   SEQUENCE   709 AA;  81271 MW;  9907236C5DB73A20 CRC64;

  Query Match             65.2%;  Score 90;  DB 13;  Length 709;
  Best Local Similarity   60.9%;  Pred. No. 8.2e-05;
  Matches   14;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||  ||| ||::|:  :||
Db    422 RNPDGDHHGPWCYTSDPKTEFDY 444


RESULT 8
Q91XG8
ID   Q91XG8      PRELIMINARY;      PRT;   716 AA.
AC   Q91XG8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010551; AAH10551.1; -.
SQ   SEQUENCE   716 AA;  80693 MW;  12474C48A7D4B46D CRC64;

  Query Match             65.2%;  Score 90;  DB 11;  Length 716;
  Best Local Similarity   65.2%;  Pred. No. 8.3e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||  ||| || :|  |:||
Db    429 RNPDGDSHGPWCYTLDPDILFDY 451


RESULT 9
Q13208
ID   Q13208      PRELIMINARY;      PRT;   567 AA.
AC   Q13208;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN HOMOLOG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20191171; PubMed=10728827;
RA   Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
RA   Carritt B.;
RT   "Structure of the human D1F15S1A locus: a chromosome 1 locus with 97%
RT   identity to the chromosome 3 gene coding for hepatocyte growth factor-
RT   like protein.";
RL   DNA Seq. 8:409-413(1998).
DR   EMBL; U28054; AAC63092.1; -.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.977; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   567 AA;  64117 MW;  3FC38B07F1645810 CRC64;

  Query Match             63.8%;  Score 88;  DB 4;  Length 567;
  Best Local Similarity   60.9%;  Pred. No. 0.00013;
  Matches   14;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          :|||||  ||| || :||  :||
Db    395 QNPDGDSHGPWCYTMDPRTPFDY 417


RESULT 10
Q9PU78
ID   Q9PU78      PRELIMINARY;      PRT;   313 AA.
AC   Q9PU78;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN (FRAGMENT).
OS   Crocodylus niloticus (Nile crocodile) (African crocodile).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX   NCBI_TaxID=8501;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20022983; PubMed=10555283;
RA   Hughes S., Zelus D., Mouchiroud D.;
RT   "Warm-blooded isochore structure in nile crocodile and turtle.";
RL   Mol. Biol. Evol. 16:1521-1527(1999).
DR   EMBL; AJ011396; CAB56422.1; -.
DR   HSSP; P00747; 1HPJ.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
FT   NON_TER     313    313
SQ   SEQUENCE   313 AA;  34793 MW;  8E084704958B5AA2 CRC64;

  Query Match             62.3%;  Score 86;  DB 13;  Length 313;
  Best Local Similarity   60.9%;  Pred. No. 0.00013;
  Matches   14;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||  |  ||| || :||  :||
Db     80 RNPDNDSHGPWCYTMDPRTPFDY 102


RESULT 11
Q90865
```

```
ID   Q90865      PRELIMINARY;      PRT;   704 AA.
AC   Q90865;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE/MACROPHAGE STIMULATING PROTEIN.
GN   HGF1/MSP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=96029010; PubMed=7554499;
RA   Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT   "Expression of HGF/SF, HGF1/MSP and c-met suggests new functions
RT   during early chick development.";
RL   Dev. Genet. 17:90-101(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; X84043; CAA58862.1; -.
DR   HSSP; P00747; 1CEA.
DR   MEROPS; S01.977; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   704 AA;  79341 MW;  CAB0D8CC41367C37 CRC64;

  Query Match             62.3%;  Score 86;  DB 13;  Length 704;
  Best Local Similarity   60.9%;  Pred. No. 0.00031;
  Matches   14;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || :||  :||
Db     416 RNPDNDSHGPWCYTMDPRTPFDY 438


RESULT 12
Q9H1V4
ID   Q9H1V4      PRELIMINARY;      PRT;   648 AA.
AC   Q9H1V4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DJ1182A14.3 (SIMILAR TO MST1 (MACROPHAGE STIMULATING 1 (HEPATOCYTE
DE   GROWTH FACTOR-LIKE))).
GN   DJ1182A14.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bird C.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AL137798; CAC17639.1; -.
DR   HSSP; P00747; 5HPG.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   648 AA;  72781 MW;  4CE077057350E463 CRC64;

  Query Match             60.9%;  Score 84;  DB 4;  Length 648;
  Best Local Similarity   76.5%;  Pred. No. 0.00056;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           |||||| |||| :||:|
Db     113 RNPDGDPGGPWCHTTDP 129


RESULT 13
Q9BH09
ID   Q9BH09      PRELIMINARY;      PRT;   728 AA.
AC   Q9BH09;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR HGF.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,
RA   Tsujimoto H.;
RT   "Molecular cloning of feline hepatocyte growth factor (HGF) cDNA.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AB046610; BAB21499.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   728 AA;  83067 MW;  8D7F4A333D1E190A CRC64;

  Query Match             58.7%;  Score 81;  DB 6;  Length 728;
```

```
  Best Local Similarity   60.9%;  Pred. No. 0.0017;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || ||   :||
Db     439 RNPDDDAHGPWCYTGNPLIPWDY 461


RESULT 14
Q9BGN9
ID   Q9BGN9      PRELIMINARY;      PRT;    75 AA.
AC   Q9BGN9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RA   Murakami S., Fujiwara C., Miyamoto Y., Takeuchi S., Takahashi S.,
RA   Okuda K.;
RT   "Expression and action of hepatocyte growth factor in bovine
RT   endometrial stromal and epithelial cells in vitro.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB056447; BAB33031.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00130; KR; 2.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
FT   NON_TER       1      1
FT   NON_TER      75     75
SQ   SEQUENCE   75 AA;  8831 MW;  829EEEFCC49701B1 CRC64;

  Query Match             57.2%;  Score 79;  DB 6;  Length 75;
  Best Local Similarity   54.5%;  Pred. No. 0.00029;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db       4 RNPRGEEGGPWCFTSNPEVRYE 25


RESULT 15
Q9N1B8
ID   Q9N1B8      PRELIMINARY;      PRT;   109 AA.
AC   Q9N1B8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ENDOMETRIUM FROM DAY 1 CYCLIC UTERUS;
RX   MEDLINE=20297031; PubMed=10819791;
RA   Chen C., Spencer T.E., Bazer F.W.;
RT   "Expression of hepatocyte growth factor and its receptor c-met in the
RT   ovine uterus.";
RL   Biol. Reprod. 62:1844-1850(2000).
DR   EMBL; AF213397; AAF25945.1; -.
DR   HSSP; P14210; 2HGF.
DR   InterPro; IPR000001; Kringle.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER       1      1
FT   NON_TER     109    109
SQ   SEQUENCE   109 AA;  12501 MW;  1F88FE5DBC0D4A5D CRC64;

  Query Match             57.2%;  Score 79;  DB 6;  Length 109;
  Best Local Similarity   54.5%;  Pred. No. 0.00043;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db      67 RNPRGEEGGPWCFTSNPEVRYE 88


RESULT 16
Q9BYM0
ID   Q9BYM0      PRELIMINARY;      PRT;   208 AA.
AC   Q9BYM0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR (FRAGMENT).
GN   HGF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91369928; PubMed=1832556;
RA   Miyazawa K., Kitamura A., Kitamura N.;
RT   "Structural organization and the transcription initiation site of the
RT   human hepatocyte growth factor gene.";
RL   Biochemistry 30:9170-9176(1991).
DR   EMBL; M75971; AAG53459.1; -.
DR   EMBL; M75967; AAG53459.1; JOINED.
DR   EMBL; M75966; AAG53459.1; JOINED.
DR   EMBL; M75968; AAG53459.1; JOINED.
DR   EMBL; M75969; AAG53459.1; JOINED.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR003966; Prothrombin.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23931 MW;  AE9C50DE5A86B37B CRC64;

  Query Match             57.2%;  Score 79;  DB 4;  Length 208;
  Best Local Similarity   54.5%;  Pred. No. 0.00088;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db     178 RNPRGEEGGPWCFTSNPEVRYE 199


RESULT 17
Q13494
ID   Q13494      PRELIMINARY;      PRT;   210 AA.
```

```
AC   Q13494;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HGF AGONIST/ANTAGONIST.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96278713; PubMed=8662798;
RA   Cioce V., Csaky K.G., Chan A.M.L., Bottaro D.P., Taylor W.G.,
RA   Jensen R., Aaronson S.A., Rubin J.S.;
RT   "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
RT   HGF/scatter factor variant with partial agonist/antagonist activity.";
RL   J. Biol. Chem. 271:13110-13115(1996).
DR   EMBL; U46010; AAC50539.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
SQ   SEQUENCE   210 AA;  24116 MW;  94A6EE9C50DE5A86 CRC64;

  Query Match             57.2%;  Score 79;  DB 4;  Length 210;
  Best Local Similarity   54.5%;  Pred. No. 0.00088;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db     178 RNPRGEEGGPWCFTSNPEVRYE 199


RESULT 18
O55027
ID   O55027      PRELIMINARY;      PRT;   211 AA.
AC   O55027;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR NK1.
GN   HGF.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE=98147766; PubMed=9488442;
RA   Jakubczak J.L., LaRochelle W.J., Merlino G.;
RT   "NK1, a natural splice variant of hepatocyte growth factor/scatter
RT   factor, is a partial agonist in vivo.";
RL   Mol. Cell. Biol. 18:1275-1283(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   May M., LaRochelle W.J.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF042856; AAC40051.1; -.
DR   HSSP; P14210; 1BHT.
DR   MGD; MGI:96079; Hgf.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
SQ   SEQUENCE   211 AA;  24308 MW;  758BD0687A835F48 CRC64;

  Query Match             57.2%;  Score 79;  DB 11;  Length 211;
  Best Local Similarity   54.5%;  Pred. No. 0.00089;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db     179 RNPRGEEGGPWCFTSNPEVRYE 200


RESULT 19
Q02935
ID   Q02935      PRELIMINARY;      PRT;   290 AA.
AC   Q02935;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR, HEAVY CHAIN PRECURSOR.
GN   HGF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=91200041; PubMed=1826653;
RA   Miyazawa K., Kitamura A., Naka D., Kitamura N.;
RT   "An alternatively processed mRNA generated from human hepatocyte
RT   growth factor gene.";
RL   Eur. J. Biochem. 197:15-22(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93087571; PubMed=1280830;
RA   Hartmann G., Naldini L., Weidner K.M., Sachs M., Vigna E.,
RA   Comoglio P.M., Birdmeier W.;
RT   "A functional domain in the heavy chain of scatter factor/hepatocyte
RT   growth factor binds the c-Met receptor and induces cell dissociation
RT   but not mitogenesis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS AS
CC       GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES. IT
CC       HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SIMILARITY: CONTAINS TWO KRINGLE REGIONS.
CC   -!- SIMILARITY: HIGH, TO OTHER HGF; LOWER, TO PLASMINOGEN.
DR   EMBL; X57574; CAA40802.1; -.
DR   EMBL; L02931; AAA52649.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
KW   Growth factor; Kringle; Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    290       HEPATOCYTE GROWTH FACTOR, HEAVY CHAIN.
FT   DOMAIN      128    206       KRINGLE 1 (BY SIMILARITY).
FT   DOMAIN      211    288       KRINGLE 2 (BY SIMILARITY).
FT   MOD_RES      32     32       PYRROLIDONE CARBOXYLIC ACID
FT                                (BY SIMILARITY).
```

```
SQ   SEQUENCE   290 AA;  33765 MW;  C8A18A6F0D63200A CRC64;

  Query Match             57.2%;  Score 79;  DB 4;  Length 290;
  Best Local Similarity   54.5%;  Pred. No. 0.0013;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db     178 RNPRGEEGGPWCFTSNPEVRYE 199


RESULT 20
Q14519
ID   Q14519      PRELIMINARY;      PRT;   296 AA.
AC   Q14519;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COMPETITIVE HGF ANTAGONIST.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chedid M.,
RA   Aaronson S.A.;
RT   "Identification of a Competitive HGF Antagonist Encoded by an
RT   Alternative Transcript.";
RL   Science 0:0-0(1991).
DR   EMBL; M77227; AAA35980.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00024; PAN; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
SQ   SEQUENCE   296 AA;  34546 MW;  A45E456B87AE03BE CRC64;

  Query Match             57.2%;  Score 79;  DB 4;  Length 296;
  Best Local Similarity   54.5%;  Pred. No. 0.0013;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYD 22
           ||| |: |||| :|:||   |:
Db     178 RNPRGEEGGPWCFTSNPEVRYE 199


RESULT 21
Q91691
ID   Q91691      PRELIMINARY;      PRT;   716 AA.
AC   Q91691;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GROWTH FACTOR LIVERTINE.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ruiz i Altaba A., Thery C.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57455; AAB52574.1; -.
DR   HSSP; P00747; 1CEA.
DR   MEROPS; S01.UNA; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   716 AA;  81971 MW;  508376A0E4398798 CRC64;

  Query Match             57.2%;  Score 79;  DB 13;  Length 716;
  Best Local Similarity   56.5%;  Pred. No. 0.0034;
  Matches   13;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || :|   :||
Db     426 RNPDRDSHGPWCYTMDPNTPFDY 448


RESULT 22
P70006
ID   P70006      PRELIMINARY;      PRT;   717 AA.
AC   P70006;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE=96404125; PubMed=8808403;
RA   Aberger F., Schmidt G., Richter K.;
RT   "The Xenopus homologue of hepatocyte growth factor-like protein is
RT   specifically expressed in the presumptive neural plate during
RT   gastrulation.";
RL   Mech. Dev. 54:23-37(1996).
DR   EMBL; Y08734; CAA69989.1; -.
DR   HSSP; P00747; 1CEA.
DR   MEROPS; S01.977; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 3.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    717       HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
SQ   SEQUENCE   717 AA;  82017 MW;  6F877A432C8CDD54 CRC64;

  Query Match             57.2%;  Score 79;  DB 13;  Length 717;
  Best Local Similarity   56.5%;  Pred. No. 0.0034;
```

```
  Matches   13;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||  |  ||| || :|   :||
Db     427 RNPDRDSHGPWCYTMDPNTPFDY 449


RESULT 23
Q25101
ID   Q25101      PRELIMINARY;      PRT;   385 AA.
AC   Q25101;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SERINE PROTEINASE.
GN   HMSERP1.
OS   Herdmania momus.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC   Stolidobranchia; Pyuridae; Herdmania.
OX   NCBI_TaxID=7733;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CURVATA;
RA   Arnold J.M., Kennett C., Lavin M.F.;
RT   "Transient expression of a novel serine protease in the ectoderm of
RT   the ascidian Herdmania momus during development.";
RL   Dev. Genes Evol. 206:455-463(1997).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; U63517; AAB66650.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.UPA; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   385 AA;  42935 MW;  BFB1D05D5232E6A0 CRC64;

  Query Match             56.5%;  Score 78;  DB 5;  Length 385;
  Best Local Similarity   60.9%;  Pred. No. 0.0024;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||  |  ||| ||||    :||
Db      75 RNPDQDWRGPWCYTTNEFMRWDY 97


RESULT 24
Q91402
ID   Q91402      PRELIMINARY;      PRT;   710 AA.
AC   Q91402;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR.
GN   HGF.
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
OX   NCBI_TaxID=8353;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TAILBUD;
RX   MEDLINE=95267690; PubMed=7748783;
RA   Nakamura H., Tashiro K., Nakamura T., Shiokawa K.;
RT   "Molecular cloning of Xenopus HGF cDNA and its expression studies in
RT   Xenopus early embryogenesis.";
RL   Mech. Dev. 49:123-131(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; S77422; AAB34354.2; -.
DR   HSSP; P14210; 1BHT.
DR   MEROPS; S01.976; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 3.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   710 AA;  81487 MW;  5FE6480BE31C27FC CRC64;

  Query Match             56.5%;  Score 78;  DB 13;  Length 710;
  Best Local Similarity   56.5%;  Pred. No. 0.0047;
  Matches   13;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||  |  ||| || :|   :||
Db     425 RNPDNDAHGPWCYTDDPFVPWDY 447


RESULT 25
Q90978
ID   Q90978      PRELIMINARY;      PRT;   726 AA.
AC   Q90978; Q90866;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR / SCATTER FACTOR PRECURSOR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE=96029010; PubMed=7554499;
RA   Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT   "Expression of HGF/SF, HGF1/MSP, and c-met suggests new functions
RT   during early chick development.";
RL   Dev. Genet. 17:90-101(1995).
RN   [2]
RP   SEQUENCE OF 1-409 FROM N.A.
RC   STRAIN=WHITE LEGHORN, AND RHODE ISLAND RED X LIGHT SUSSEX;
RC   TISSUE=EMBRYO;
RX   MEDLINE=95237013; PubMed=7720585;
RA   Streit A., Stern C.D., Thery C., Ireland G.W., Aparicio S.,
RA   Sharpe M.J., Gherardi E.;
RT   "A role for HGF/SF in neural induction and its expression in Hensen's
RT   node during gastrulation.";
RL   Development 121:813-824(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
```

```
DR   EMBL; X84045; CAA58864.1; -.
DR   EMBL; X80131; CAA56430.1; -.
DR   HSSP; P14210; 1BHT.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Alternative splicing; Glycoprotein; Growth factor; Hydrolase; Kringle;
KW   Serine protease; Serine protease homolog; Signal.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    726       HEPATOCYTE GROWTH FACTOR/ SCATTER FACTOR.
SQ   SEQUENCE   726 AA;  82913 MW;  5805F048A5766C38 CRC64;

  Query Match             55.1%;  Score 76;  DB 13;  Length 726;
  Best Local Similarity   56.5%;  Pred. No. 0.0094;
  Matches   13;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || :|   :||
Db     439 RNPDDDFHGPWCYTDDPLIPWDY 461


RESULT 26
Q28398
ID   Q28398      PRELIMINARY;      PRT;  2869 AA.
AC   Q28398;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
OS   Erinaceus europaeus (Western European hedgehog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
OX   NCBI_TaxID=9365;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=96025778; PubMed=7592597;
RA   Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA   Byrne C.D., Fong K.J., Meer K., Patthy L.;
RT   "The recurring evolution of lipoprotein(a). Insights from cloning of
RT   hedgehog apolipoprotein(a).";
RL   J. Biol. Chem. 270:24004-24009(1995).
DR   EMBL; U33170; AAC48522.1; -.
DR   HSSP; P00747; 1PMK.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 31.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 31.
DR   PROSITE; PS00021; KRINGLE_1; 17.
DR   PROSITE; PS50070; KRINGLE_2; 31.
KW   Lipoprotein.
FT   NON_TER       1       1
SQ   SEQUENCE   2869 AA;  318601 MW;  9527CEF985A4FB2A CRC64;

  Query Match             52.5%;  Score 72.5;  DB 6;  Length 2869;
  Best Local Similarity   56.5%;  Pred. No. 0.14;
  Matches   13;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||:|  || ||||    ::|

Db    2548 RNPDGEV-APWCYTTNSAMRWEY 2569


RESULT 27
O42341
ID   O42341      PRELIMINARY;      PRT;  215 AA.
AC   O42341;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HGF ALPHA-CHAIN (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   Isokawa K., Yahagi N., Honma J., Tanaka M., Murakami K., Yokoyama H.,
RA   Toda Y.;
RT   "The expression of hepatocyte growth factor mRNA is temporally
RT   relevant to cardiac endothelial-mesenchymal transformation.";
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D63779; BAA23643.1; -.
DR   HSSP; P00747; 1KRN.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 3.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 2.
DR   PROSITE; PS00021; KRINGLE_1; 3.
DR   PROSITE; PS50070; KRINGLE_2; 3.
FT   NON_TER       1       1
FT   NON_TER     215     215
SQ   SEQUENCE   215 AA;  24949 MW;  55E67AB52DAAF316 CRC64;

  Query Match             49.3%;  Score 68;  DB 13;  Length 215;
  Best Local Similarity   64.7%;  Pred. No. 0.037;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           |||| |  ||| || :|
Db     198 RNPDDDSHGPWCYTDDP 214


RESULT 28
Q16609
ID   Q16609      PRELIMINARY;      PRT;  132 AA.
AC   Q16609;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   (APOARGC).
GN   APOARGC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=95268939; PubMed=7749817;
RA   Byrne C.D., Schwartz K., Lawn R.M.;
RT   "Loss of a splice donor site at a 'skipped exon' in a gene homologous
RT   to apolipoprotein(a) leads to an mRNA encoding a protein consisting of
RT   a single kringle domain.";
RL   Arterioscler. Thromb. Vasc. Biol. 15:65-70(1995).
DR   EMBL; U19518; AAA85693.1; -.
DR   EMBL; U19517; AAA85692.1; -.
DR   HSSP; P00747; 1PMK.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
```

```
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
SQ   SEQUENCE   132 AA;  14886 MW;  3794AD30A586DBBA CRC64;

  Query Match             48.2%;  Score 66.5;  DB 4;  Length 132;
  Best Local Similarity   52.2%;  Pred. No. 0.036;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||    ||| |||:|   ::|
Db      78 RNPDCS-AGPWCYTTDPNVRWEY 99


RESULT 29
Q90WS2
ID   Q90WS2      PRELIMINARY;      PRT;   385 AA.
AC   Q90WS2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE THROMBIN (FRAGMENT).
GN   THROMBIN.
OS   Elaphe sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Colubridae; Colubrinae; Elaphe.
OX   NCBI_TaxID=114965;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Hughes S., Laurin M., Gouy M., Mouchiroud D.;
RT   "The phylogenetic position of turtles among sauropsidae remains
RT   contentious with new nuclear data.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ286868; CAC69543.1; -.
FT   NON_TER       1      1
FT   NON_TER     385    385
SQ   SEQUENCE   385 AA;  43901 MW;  BF9E59B93C0C8B7B CRC64;

  Query Match             47.8%;  Score 66;  DB 13;  Length 385;
  Best Local Similarity   47.8%;  Pred. No. 0.14;
  Matches   11;  Conservative    2;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  | | |  :|   |:|
Db      91 RNPDHDEEGVWCYVDHPNMTYNY 113


RESULT 30
O46506
ID   O46506      PRELIMINARY;      PRT;   454 AA.
AC   O46506;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN A (FRAGMENT).
GN   BABAPOA.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cox L.A., Jett C., Hixson J.E.;
RT   "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT   Site Mutation is Associated with Deletion of a Single Exon in a Null
RT   Allele.";
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF029691; AAB97886.1; -.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.226; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Lipoprotein; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   454 AA;  50041 MW;  974E30744C187B2F CRC64;

  Query Match             46.7%;  Score 64.5;  DB 6;  Length 454;
  Best Local Similarity   47.8%;  Pred. No. 0.27;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| :| :|   ::|
Db     152 RNPDADT-GPWCFTMDPSVRWEY 173


RESULT 31
Q91001
ID   Q91001      PRELIMINARY;      PRT;   607 AA.
AC   Q91001;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   THROMBIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., MacGillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=94223694; PubMed=7513365;
RA   Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;
RT   "Evolution of prothrombin: isolation and characterization of the cDNAs
RT   encoding chicken and hagfish prothrombin.";
RL   J. Mol. Evol. 38:177-187(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Banfield D.K.;
RL   Submitted (DEC-1991) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M81391; AAA21619.1; -.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
```

```
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   607 AA;  69110 MW;  002F3606EA36270F CRC64;

  Query Match             46.4%;  Score 64;  DB 13;  Length 607;
  Best Local Similarity   58.8%;  Pred. No. 0.44;
  Matches   10;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNP 17
          ||||  :  ||| || :|
Db    157 RNPDNNSEGPWCYTRDP 173


RESULT 32
Q9PTW7
ID   Q9PTW7      PRELIMINARY;      PRT;   608 AA.
AC   Q9PTW7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROTHROMBIN.
GN   OSPT.
OS   Struthio camelus (Ostrich).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC   Struthio.
OX   NCBI_TaxID=8801;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20579470; PubMed=11137455;
RA   Frost C., Naude R., Oelofsen W., Muramoto K., Naganuma T., Ogawa T.;
RT   "Purification and characterization of ostrich prothrombin.";
RL   Int. J. Biochem. Cell Biol. 32:1151-1159(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AB028871; BAA89046.1; -.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   608 AA;  69392 MW;  11B974B9AEE54EA2 CRC64;

  Query Match             46.4%;  Score 64;  DB 13;  Length 608;
  Best Local Similarity   58.8%;  Pred. No. 0.44;
  Matches   10;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNP 17
          ||||  :  ||| || :|
Db    157 RNPDNNPEGPWCYTRDP 173


RESULT 33
Q9UIR5
ID   Q9UIR5      PRELIMINARY;      PRT;   113 AA.
AC   Q9UIR5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
GN   APOA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181705; PubMed=11285247;
RA   Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT   "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT   types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT   different patterns in Africans and Caucasians.";
RL   Hum. Mol. Genet. 10:815-824(2001).
DR   EMBL; AF158663; AAF03680.1; -.
DR   EMBL; AF158662; AAF03680.1; JOINED.
DR   HSSP; P00747; 1PMK.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     113    113
SQ   SEQUENCE   113 AA;  12685 MW;  F3D65681D9B5253A CRC64;

  Query Match             46.0%;  Score 63.5;  DB 4;  Length 113;
  Best Local Similarity   47.8%;  Pred. No. 0.083;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||  |  |||  :|  :|   ::|
Db     61 RNPDADT-GPWCFTMDPSIRWEY 82


RESULT 34
Q90WP0
ID   Q90WP0      PRELIMINARY;      PRT;   378 AA.
AC   Q90WP0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE THROMBIN (FRAGMENT).
GN   THROMBIN.
OS   Trachemys scripta elegans.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
OX   NCBI_TaxID=31138;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Hughes S., Laurin M., Gouy M., Mouchiroud D.;
```

```
RT   "The phylogenetic position of turtles among sauropsidae remains
RT   contentious with new nuclear data.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ286869; CAC69549.1; -.
FT   NON_TER       1      1
FT   NON_TER     378    378
SQ   SEQUENCE   378 AA;  42817 MW;  62F09CC7A81059DC CRC64;

  Query Match             45.7%;  Score 63;  DB 13;  Length 378;
  Best Local Similarity   43.5%;  Pred. No. 0.37;
  Matches   10;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  | | |  :|   ::|
Db      88 RNPDDDDEGVWCYVDHPNTTFEY 110


RESULT 35
Q9UIR8
ID   Q9UIR8      PRELIMINARY;      PRT;   105 AA.
AC   Q9UIR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
GN   APOA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181705; PubMed=11285247;
RA   Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT   "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT   types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT   different patterns in Africans and Caucasians.";
RL   Hum. Mol. Genet. 10:815-824(2001).
DR   EMBL; AF158656; AAF03677.1; -.
DR   EMBL; AF158655; AAF03677.1; JOINED.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     105    105
SQ   SEQUENCE   105 AA;  11882 MW;  6ECB6C02CD30EFA2 CRC64;

  Query Match             45.3%;  Score 62.5;  DB 4;  Length 105;
  Best Local Similarity   43.5%;  Pred. No. 0.11;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| ::  || || :|   ::|
Db      61 RNPDAEI-SPWCYTMDPNVRWEY 82


RESULT 36
O77688
ID   O77688      PRELIMINARY;      PRT;   111 AA.
AC   O77688;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROTHROMBIN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=LIVER;
RX   MEDLINE=99003227; PubMed=9786880;
RA   Lee T.H., Rhim T., Kim S.S.;
RT   "Prothrombin kringle-2 domain has a growth inhibitory activity against
RT   basic fibroblast growth factor-stimulated capillary endothelial
RT   cells.";
RL   J. Biol. Chem. 273:28805-28812(1998).
DR   EMBL; AF080065; AAC71006.1; -.
DR   HSSP; P00735; 1A0H.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER       1      1
FT   NON_TER     111    111
SQ   SEQUENCE   111 AA;  12168 MW;  CB86624D8D7F9DE9 CRC64;

  Query Match             43.5%;  Score 60;  DB 6;  Length 111;
  Best Local Similarity   43.5%;  Pred. No. 0.27;
  Matches   10;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  | | |       ::|
Db      65 RNPDGDEEGAWRYVAGRPSSFEY 87


RESULT 37
Q13209
ID   Q13209      PRELIMINARY;      PRT;   286 AA.
AC   Q13209;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN HOMOLOG (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20191171; PubMed=10728827;
RA   Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
RA   Carritt B.;
RT   "Structure of the human D1F15S1A locus: a chromosome 1 locus with 97%
RT   identity to the chromosome 3 gene coding for hepatocyte growth factor-
RT   like protein.";
RL   DNA Seq. 8:409-413(1998).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; U28055; AAC35412.1; -.
DR   HSSP; P00747; 5HPG.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Hydrolase; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   286 AA;  31986 MW;  43489B747C8D3F90 CRC64;

  Query Match             43.5%;  Score 60;  DB 4;  Length 286;
  Best Local Similarity   60.0%;  Pred. No. 0.75;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      9 GPWAYTTNPRKLYDY 23
          ||| || :||  :||
Db      3 GPWCYTMDPRTPFDY 17


RESULT 38
Q9UKJ7
ID   Q9UKJ7      PRELIMINARY;      PRT;    60 AA.
AC   Q9UKJ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
GN   APOA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181705; PubMed=11285247;
RA   Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT   "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT   types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT   different patterns in Africans and Caucasians.";
RL   Hum. Mol. Genet. 10:815-824(2001).
DR   EMBL; AF158657; AAF03676.1; -.
DR   HSSP; P00747; 1KRN.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER      60     60
SQ   SEQUENCE   60 AA;  6799 MW;  5719AA26B3E0FF1D CRC64;

  Query Match             43.1%;  Score 59.5;  DB 4;  Length 60;
  Best Local Similarity   43.5%;  Pred. No. 0.16;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||| ::  || || :|   ::|
Db      8 RNPDAEI-RPWCYTMDPSVRWEY 29


RESULT 39
Q9UIR7
ID   Q9UIR7      PRELIMINARY;      PRT;   113 AA.
AC   Q9UIR7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
GN   APOA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181705; PubMed=11285247;
RA   Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT   "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT   types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT   different patterns in Africans and Caucasians.";
RL   Hum. Mol. Genet. 10:815-824(2001).
DR   EMBL; AF158659; AAF03678.1; -.
DR   EMBL; AF158658; AAF03678.1; JOINED.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     113    113
SQ   SEQUENCE   113 AA;  12815 MW;  4F80ADF8708548CB CRC64;

  Query Match             43.1%;  Score 59.5;  DB 4;  Length 113;
  Best Local Similarity   43.5%;  Pred. No. 0.32;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          |||| ::  || || :|   ::|
Db     61 RNPDAEI-RPWCYTMDPSVRWEY 82


RESULT 40
Q90504
ID   Q90504      PRELIMINARY;      PRT;   420 AA.
AC   Q90504;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   THROMBIN.
OS   Eptatretus stoutii (Pacific hagfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniformes;
OC   Myxinidae; Eptatretinae; Eptatretus.
OX   NCBI_TaxID=7765;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., MacGillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=94223694; PubMed=7513365;
RA   Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;
RT   "Evolution of prothrombin: isolation and characterization of the cDNAs
RT   encoding chicken and hagfish prothrombin.";
RL   J. Mol. Evol. 38:177-187(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Banfield D.K.;
RL   Submitted (DEC-1991) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; M81393; AAA21620.1; -.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   420 AA;  47888 MW;  64522AA21A57B67A CRC64;
```

```
  Query Match             42.8%;  Score 59;  DB 13;  Length 420;
  Best Local Similarity   71.4%;  Pred. No. 1.6;
  Matches   10;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYT 14
          ||||||  | | ||
Db     64 RNPDGDSEGVWCYT 77


RESULT 41
O00318
ID   O00318      PRELIMINARY;      PRT;   263 AA.
AC   O00318;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   WUGSC:DJ515N1.2 PROTEIN.
GN   WUGSC:DJ515N1.2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Du Z., Scheet P., Harper M.;
RT   "The sequence of H. sapiens PAC clone RP3-515N1.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC002073; AAB54054.1; -.
DR   HSSP; P00749; 1KDU.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
SQ   SEQUENCE   263 AA;  28248 MW;  197C3EEE8E54A242 CRC64;

  Query Match             40.6%;  Score 56;  DB 4;  Length 263;
  Best Local Similarity   60.0%;  Pred. No. 2.6;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTT 15
          ||||  |  ||| | :
Db     71 RNPDEDPRGPWCYVS 85


RESULT 42
Q96FE7
ID   Q96FE7      PRELIMINARY;      PRT;   263 AA.
AC   Q96FE7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   UNKNOWN (PROTEIN FOR MGC:17330).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND ANAPLASTIC OLIGODENDROGLIOMA WITH 1P/19Q LOSS;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011049; AAH11049.1; -.
SQ   SEQUENCE   263 AA;  28234 MW;  197C3EEE888FA242 CRC64;

  Query Match             40.6%;  Score 56;  DB 4;  Length 263;
  Best Local Similarity   60.0%;  Pred. No. 2.6;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTT 15
          ||||  |  ||| | :
Db     71 RNPDEDPRGPWCYVS 85


RESULT 43
Q91VP2
ID   Q91VP2      PRELIMINARY;      PRT;   559 AA.
AC   Q91VP2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO PLASMINOGEN ACTIVATOR, TISSUE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BREAST TUMOR;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011256; AAH11256.1; -.
SQ   SEQUENCE   559 AA;  63122 MW;  8CCEE2BDB94514D9 CRC64;

  Query Match             40.6%;  Score 56;  DB 11;  Length 559;
  Best Local Similarity   50.0%;  Pred. No. 5.9;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    2;  Gaps    2;

Qy      1 RNPDGDVGGPWAYTTNPRKL-YDY 23
          ||||||   || :    ||| ::|
Db    266 RNPDGD-ARPWCHVMKDRKLTWEY 288


RESULT 44
Q9UIR6
ID   Q9UIR6      PRELIMINARY;      PRT;   113 AA.
AC   Q9UIR6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   APOLIPOPROTEIN(A) (FRAGMENT).
GN   APOA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181705; PubMed=11285247;
RA   Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT   "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT   types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT   different patterns in Africans and Caucasians.";
RL   Hum. Mol. Genet. 10:815-824(2001).
DR   EMBL; AF158661; AAF03679.1; -.
DR   EMBL; AF158660; AAF03679.1; JOINED.
DR   HSSP; P00747; 2PK4.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     113    113
SQ   SEQUENCE   113 AA;  12697 MW;  51D4461D9C66312E CRC64;

  Query Match             40.2%;  Score 55.5;  DB 4;  Length 113;
  Best Local Similarity   47.8%;  Pred. No. 1.2;
```

```
  Matches   11;  Conservative    3;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||     || |||:|   ::|
Db      61 RNPDSG-KQPWCYTTDPCVRWEY 82


RESULT 45
Q90WT4
ID   Q90WT4      PRELIMINARY;      PRT;   382 AA.
AC   Q90WT4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE THROMBIN (FRAGMENT).
GN   THROMBIN.
OS   Crocodylus niloticus (Nile crocodile) (African crocodile).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX   NCBI_TaxID=8501;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Hughes S., Laurin M., Gouy M., Mouchiroud D.;
RT   "The phylogenetic position of turtles among sauropsidae remains
RT   contentious with new nuclear data.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ286873; CAC69536.1; -.
FT   NON_TER       1       1
FT   NON_TER     382     382
SQ   SEQUENCE   382 AA;  43282 MW;  FE8E60239D8CBECB CRC64;

  Query Match             39.9%;  Score 55;  DB 13;  Length 382;
  Best Local Similarity   43.5%;  Pred. No. 5.5;
  Matches   10;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  | |  |      ::|
Db      89 RNPDGDEEGVWCVTQGEPDNFEY 111


RESULT 46
Q9UGS5
ID   Q9UGS5      PRELIMINARY;      PRT;   117 AA.
AC   Q9UGS5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BK747E2.4 (NOVEL KRINGLE AND CUB DOMAIN PROTEIN) (FRAGMENT).
GN   BK57G9.1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ho S.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL021393; CAB62959.1; -.
DR   HSSP; P00747; 1PML.
DR   InterPro; IPR000001; Kringle.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
FT   NON_TER       1       1
FT   NON_TER     117     117
SQ   SEQUENCE   117 AA;  12673 MW;  F4270BBA1D03EDCE CRC64;

  Query Match             39.5%;  Score 54.5;  DB 4;  Length 117;
  Best Local Similarity   76.9%;  Pred. No. 1.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAY 13
           |||||||  || |
Db      87 RNPDGDV-SPWCY 98


RESULT 47
Q90Y90
ID   Q90Y90      PRELIMINARY;      PRT;   452 AA.
AC   Q90Y90;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KREMEN.
GN   KREMEN.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21167372; PubMed=11267660;
RA   Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA   Matsumoto K., Nakamura T.;
RT   "Molecular cloning and characterization of Kremen, a novel kringle-
RT   containing transmembrane protein.";
RL   Biochim. Biophys. Acta 1518:63-72(2001).
DR   EMBL; AB070851; BAB64294.1; -.
SQ   SEQUENCE   452 AA;  50188 MW;  ED24BCD1AF4564E2 CRC64;

  Query Match             39.5%;  Score 54.5;  DB 13;  Length 452;
  Best Local Similarity   76.9%;  Pred. No. 7.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAY 13
           |||||||  || |
Db      83 RNPDGDV-SPWCY 94


RESULT 48
Q9BY70
ID   Q9BY70      PRELIMINARY;      PRT;   473 AA.
AC   Q9BY70;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KRINGLE-CONTAINING TRANSMEMBRANE PROTEIN.
GN   KREMEN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21167372; PubMed=11267660;
RA   Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA   Matsumoto K., Nakamura T.;
RT   "Molecular cloning and characterization of Kremen, a novel kringle-
RT   containing transmembrane protein.";
RL   Biochim. Biophys. Acta 1518:63-72(2001).
CC   -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR   EMBL; AB059618; BAB40969.1; -.
DR   HSSP; P00747; 1PML.
DR   InterPro; IPR000859; CUB.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR002889; WSC.
DR   Pfam; PF00431; CUB; 1.
DR   Pfam; PF01822; WSC; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00042; CUB; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00321; WSC; 1.
```

```
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Transmembrane.
SQ   SEQUENCE   473 AA;  51744 MW;  F6D30DDE708C186B CRC64;

  Query Match             39.5%;  Score 54.5;  DB 4;  Length 473;
  Best Local Similarity   76.9%;  Pred. No. 8.2;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAY 13
          |||||||  || |
Db     85 RNPDGDV-SPWCY 96


RESULT 49
Q99N43
ID   Q99N43      PRELIMINARY;      PRT;   473 AA.
AC   Q99N43;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KRINGLE-CONTAINING TRANSMEMBRANE PROTEIN.
GN   KREMEN OR KREMEN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21167372; PubMed=11267660;
RA   Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA   Matsumoto K., Nakamura T.;
RT   "Molecular cloning and characterization of Kremen, a novel kringle-
RT   containing transmembrane protein.";
RL   Biochim. Biophys. Acta 1518:63-72(2001).
CC   -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR   EMBL; AB059617; BAB40968.1; -.
DR   HSSP; P00747; 1CEA.
DR   MGD; MGI:1933988; Kremen.
DR   InterPro; IPR000859; CUB.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR002889; WSC.
DR   Pfam; PF00431; CUB; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF01822; WSC; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00042; CUB; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00321; WSC; 1.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Transmembrane.
SQ   SEQUENCE   473 AA;  51716 MW;  586827788BE3FDD1 CRC64;

  Query Match             39.5%;  Score 54.5;  DB 11;  Length 473;
  Best Local Similarity   76.9%;  Pred. No. 8.2;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAY 13
          |||||||  || |
Db     85 RNPDGDV-SPWCY 96


RESULT 50
Q924S4
ID   Q924S4      PRELIMINARY;      PRT;   473 AA.
AC   Q924S4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KRINGLE-CONTAINING TRANSMEMBRANE PROTEIN.
GN   KREMEN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21167372; PubMed=11267660;
RA   Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,
RA   Nakamura T.;
RT   "Molecular cloning and characterization of Kremen, a novel kringle-
RT   containing transmembrane protein.";
RL   Biochim. Biophys. Acta 1518:63-72(2001).
DR   EMBL; AB065090; BAB62003.1; -.
KW   Transmembrane.
SQ   SEQUENCE   473 AA;  51869 MW;  9B510857DF856F08 CRC64;

  Query Match             39.5%;  Score 54.5;  DB 11;  Length 473;
  Best Local Similarity   76.9%;  Pred. No. 8.2;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      1 RNPDGDVGGPWAY 13
          |||||||  || |
Db     85 RNPDGDV-SPWCY 96


Search completed: November  8, 2002, 09:33:33
Job time : 29 secs
```

(9)

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 8, 2002, 09:31:20 ; Search time 10 Seconds
(without alignments)
89.055 Million cell updates/sec

Title: US-09-657-431-9
Perfect score: 138
Sequence: 1 RNPDGDVGGPWAYTTNPRKLYDY 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	97.1	810	1	PLMN_HUMAN	P00747 homo sapien
2	130	94.2	810	1	PLMN_MACMU	P12545 macaca mula
3	128	92.8	812	1	PLMN_MOUSE	P20918 mus musculu
4	124	89.9	343	1	PLMN_SHEEP	P81286 ovis aries
5	120	87.0	810	1	PLMN_ERIEU	Q29485 erinaceus e
6	117	84.8	4548	1	APOA_HUMAN	P08519 homo sapien
7	114	82.6	338	1	PLMN_HORSE	P80010 equus cabal
8	113	81.9	790	1	PLMN_PIG	P06867 sus scrofa
9	111	80.4	812	1	PLMN_BOVIN	P06868 bos taurus
10	110	79.7	333	1	PLMN_CANFA	P80009 canis famil
11	92	66.7	711	1	HGFL_HUMAN	P26927 homo sapien
12	90	65.2	716	1	HGFL_MOUSE	P26928 mus musculu
13	82	59.4	728	1	HGF_RAT	P17945 rattus norv
14	81	58.7	728	1	HGF_HUMAN	P14210 homo sapien
15	81	58.7	728	1	HGF_MOUSE	Q08048 mus musculu
16	80	58.0	625	1	THRB_BOVIN	P00735 bos taurus
17	77	55.8	169	1	PLMN_RAT	Q01177 rattus norv
18	73	52.9	325	1	PLMN_PETMA	P33574 petromyzon
19	73	52.9	617	1	THRB_RAT	P18292 rattus norv
20	71	51.4	618	1	THRB_MOUSE	P19221 mus musculu
21	71	51.4	622	1	THRB_HUMAN	P00734 homo sapien
22	65.5	47.5	1420	1	APOA_MACMU	P14417 macaca mula
23	60	43.5	435	1	YE90_MYCTU	P71771 mycobacteri
24	59	42.8	943	1	ROR2_HUMAN	Q01974 homo sapien
25	59	42.8	944	1	ROR2_MOUSE	Q9z138 mus musculu
26	56	40.6	559	1	TPA_MOUSE	P11214 mus musculu
27	56	40.6	559	1	TPA_RAT	P19637 rattus norv
28	52.5	38.0	700	1	PHLN_BURPS	Q9rgs8 burkholderi
29	52	37.7	562	1	TPA_HUMAN	P00750 homo sapien
30	51	37.0	194	1	Y4HQ_RHISN	P50361 rhizobium s
31	51	37.0	198	1	EST1_SCHGA	P81429 schizaphis
32	51	37.0	552	1	ESTE_MYZPE	P35501 myzus persi
33	51	37.0	564	1	ESTF_MYZPE	P35502 myzus persi
34	51	37.0	566	1	TPA_BOVIN	Q28198 bos taurus
35	50.5	36.6	394	1	URTG_DESRO	P49150 desmodus ro
36	49	35.5	615	1	FA12_HUMAN	P00748 homo sapien
37	48.5	35.1	961	1	PMA7_ARATH	Q9ly32 arabidopsis
38	48	34.8	268	1	Y145_METJA	Q57609 methanococc
39	48	34.8	517	1	PHLC_MYCTU	P95245 mycobacteri
40	47	34.1	520	1	GAG_SIVA1	P27972 simian immu
41	46.5	33.7	129	1	TRD5_ECOLI	P27192 escherichia
42	46	33.3	347	1	POR2_HUMAN	P45880 homo sapien
43	46	33.3	519	1	GAG_SIVAT	P05892 simian immu
44	46	33.3	521	1	GAG_SIVAG	P27978 simian immu
45	46	33.3	603	1	FA12_CAVPO	Q04962 cavia porce
46	45.5	33.0	373	1	YHX6_YEAST	P38866 saccharomyc
47	45.5	33.0	514	1	OASL_HUMAN	Q15646 homo sapien
48	45.5	33.0	607	1	PPO_VITVI	P43311 vitis vinif
49	45.5	33.0	761	1	NETR_MOUSE	O08762 mus musculu
50	45	32.6	427	1	FOLD_YEAST	Q12676 saccharomyc

ALIGNMENTS

RESULT 1
PLMN_HUMAN
ID PLMN_HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,

RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vali Z., Patthy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vali Z., Patthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RT are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RT plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE 268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevich R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;
RL Submitted (JUN-1995) to the PDB data bank.
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";

```
RL   J. Mol. Biol. 212:541-552(1990).
RN   [27]
RP   VARIANTS PHE-374 AND THR-620.
RX   MEDLINE=91095410; PubMed=1986355;
RA   Ichinose A., Espling E.S., Takamatsu J., Saito H., Shinmyozu K.,

  Query Match             97.1%;  Score 134;  DB 1;  Length 810;
  Best Local Similarity   95.7%;  Pred. No. 1e-11;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||||||| |||||||||||
Db     532 RNPDGDVGGPWCYTTNPRKLYDY 554


RESULT 2
PLMN_MACMU
ID   PLMN_MACMU     STANDARD;      PRT;   810 AA.
AC   P12545;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89174660; PubMed=2925643;
RA   Tomlinson J.E., McLean J.W., Lawn R.M.;
RT   "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT   synthesis.";
RL   J. Biol. Chem. 264:5957-5965(1989).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-580, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J04697; AAA36901.1; -.
DR   PIR; B30848; B30848.
DR   PIR; B32869; B32869.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL       1    19
FT   CHAIN       20   810       PLASMINOGEN.
FT   CHAIN       20   580       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE     20    96       ACTIVATION PEPTIDE.
FT   CHAIN       97   580       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN      581   810       PLASMIN LIGHT CHAIN B.
FT   DOMAIN     103   181       KRINGLE 1.
FT   DOMAIN     184   262       KRINGLE 2.
FT   DOMAIN     275   352       KRINGLE 3.
FT   DOMAIN     377   454       KRINGLE 4.
FT   DOMAIN     481   560       KRINGLE 5.
FT   DOMAIN     581   810       SERINE PROTEASE.
FT   ACT_SITE   622   622       CHARGE RELAY SYSTEM.
FT   ACT_SITE   665   665       CHARGE RELAY SYSTEM.
FT   ACT_SITE   760   760       CHARGE RELAY SYSTEM.
FT   BINDING    136   136       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    158   158       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    172   172       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    432   432       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    445   445       OMEGA-AMINOCARBOXYLIC ACIDS.
FT   BINDING    134   134       FIBRIN.
FT   BINDING    136   136       FIBRIN.
FT   DISULFID    49    73       BY SIMILARITY.
FT   DISULFID    53    61       BY SIMILARITY.
FT   DISULFID   103   181       BY SIMILARITY.
FT   DISULFID   124   164       BY SIMILARITY.
FT   DISULFID   152   176       BY SIMILARITY.
FT   DISULFID   185   262       BY SIMILARITY.
FT   DISULFID   188   316       BY SIMILARITY.
FT   DISULFID   206   245       BY SIMILARITY.
FT   DISULFID   234   257       BY SIMILARITY.
FT   DISULFID   275   352       BY SIMILARITY.
FT   DISULFID   296   335       BY SIMILARITY.
FT   DISULFID   324   347       BY SIMILARITY.
FT   DISULFID   377   454       BY SIMILARITY.
FT   DISULFID   398   437       BY SIMILARITY.
FT   DISULFID   426   449       BY SIMILARITY.
FT   DISULFID   481   560       BY SIMILARITY.
FT   DISULFID   502   543       BY SIMILARITY.
FT   DISULFID   531   555       BY SIMILARITY.
FT   DISULFID   567   685       BY SIMILARITY.
FT   DISULFID   577   585       BY SIMILARITY.
FT   DISULFID   607   623       BY SIMILARITY.
FT   DISULFID   699   766       BY SIMILARITY.
FT   DISULFID   729   745       BY SIMILARITY.
FT   DISULFID   756   784       BY SIMILARITY.
FT   CARBOHYD   365   365       O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   810 AA;  90255 MW;  A75E1C51A1A0F24A CRC64;

  Query Match             94.2%;  Score 130;  DB 1;  Length 810;
  Best Local Similarity   91.3%;  Pred. No. 4e-11;
  Matches   21;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||||||||  ||||||||:||
Db     532 RNPDGDVGGPWCYTTNPRKLFDY 554
```

```
RESULT 3
PLMN_MOUSE
ID   PLMN_MOUSE     STANDARD;      PRT;   812 AA.
AC   P20918;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN   PLG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91184812; PubMed=2081600;
RA   Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse plasminogen and
RT   localization of the gene to mouse chromosome 17.";
RL   Genomics 8:49-61(1990).
RN   [2]
RP   CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX   MEDLINE=95042728; PubMed=7525077;
RA   O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA   Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT   "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT   suppression of metastases by a Lewis lung carcinoma.";
RL   Cell 79:315-328(1994).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC       NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC       METASTATIC TUMORS IN VIVO.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC       INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC       TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC       ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04766; AAA50168.1; -.
DR   PIR; A38514; A38514.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   MGD; MGI:97620; Plg.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL       1     19
FT   CHAIN       20    812       PLASMINOGEN.
FT   CHAIN       20    581       PLASMIN HEAVY CHAIN A.
FT   PEPTIDE     20     97       ACTIVATION PEPTIDE.
FT   CHAIN       98    581       PLASMIN SHORT FORM OF CHAIN A.
FT   CHAIN       98   ?436       ANGIOSTATIN.
FT   CHAIN      582    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN     103    181       KRINGLE 1.
FT   DOMAIN     184    262       KRINGLE 2.
FT   DOMAIN     275    352       KRINGLE 3.
FT   DOMAIN     377    454       KRINGLE 4.
FT   DOMAIN     481    560       KRINGLE 5.
FT   DOMAIN     582    812       SERINE PROTEASE.
FT   ACT_SITE   624    624       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   667    667       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   762    762       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    49     73       BY SIMILARITY.
FT   DISULFID    53     61       BY SIMILARITY.
FT   DISULFID   103    181       BY SIMILARITY.
FT   DISULFID   124    164       BY SIMILARITY.
FT   DISULFID   152    176       BY SIMILARITY.
FT   DISULFID   185    262       BY SIMILARITY.
FT   DISULFID   188    316       BY SIMILARITY.
FT   DISULFID   206    245       BY SIMILARITY.
FT   DISULFID   234    257       BY SIMILARITY.
FT   DISULFID   275    352       BY SIMILARITY.
FT   DISULFID   296    335       BY SIMILARITY.
FT   DISULFID   324    347       BY SIMILARITY.
FT   DISULFID   377    454       BY SIMILARITY.
FT   DISULFID   398    437       BY SIMILARITY.
FT   DISULFID   426    449       BY SIMILARITY.
FT   DISULFID   481    560       BY SIMILARITY.
FT   DISULFID   502    543       BY SIMILARITY.
FT   DISULFID   531    555       BY SIMILARITY.
FT   DISULFID   568    687       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   578    586       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   609    625       BY SIMILARITY.
FT   DISULFID   701    768       BY SIMILARITY.
FT   DISULFID   731    747       BY SIMILARITY.
FT   DISULFID   758    786       BY SIMILARITY.
SQ   SEQUENCE   812 AA;  90846 MW;  D34A74A4FC2256F8 CRC64;

  Query Match             92.8%;  Score 128;  DB 1;  Length 812;
  Best Local Similarity   91.3%;  Pred. No. 7.7e-11;
  Matches   21;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| ||||||||||
Db     532 RNPDGDVNGPWCYTTNPRKLYDY 554


RESULT 4
PLMN_SHEEP
ID   PLMN_SHEEP     STANDARD;      PRT;   343 AA.
AC   P81286;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
```

```
GN   PLG.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93149995; PubMed=1492092;
RA   Schaller J., Straub C., Kampfer U., Rickli E.E.;
RT   "Complete amino acid sequence of ovine miniplasminogen.";
RL   Protein Seq. Data Anal. 5:21-25(1992).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 2 KRINGLE DOMAINS.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   DOMAIN       <1    140       HEAVY CHAIN A.
FT   DOMAIN      141   >343       LIGHT CHAIN A.
FT   DOMAIN       <1     17       KRINGLE 4.
FT   DOMAIN       41    120       KRINGLE 5.
FT   DOMAIN      114    341       SERINE PROTEASE.
FT   ACT_SITE    181    181       CHARGE RELAY SYSTEM.
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM.
FT   ACT_SITE    319    319       CHARGE RELAY SYSTEM.
FT   NON_TER     343    343
SQ   SEQUENCE   343 AA;  37662 MW;  8DF6EBA92D596EE0 CRC64;

  Query Match             89.9%;  Score 124;  DB 1;  Length 343;
  Best Local Similarity   87.0%;  Pred. No. 1.2e-10;
  Matches   20;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| ||||||||:||
Db      66 RNPDGDVNGPWCYTTNPRKLFDY 88


RESULT 5
PLMN_ERIEU
ID   PLMN_ERIEU     STANDARD;      PRT;   810 AA.
AC   Q29485;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Erinaceus europaeus (Western European hedgehog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
OX   NCBI_TaxID=9365;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96025778; PubMed=7592597;
RA   Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA   Byrne C.D., Fong K.J., Meer K., Patthy L.;
RT   "The recurring evolution of lipoprotein(a). Insights from cloning of
RT   hedgehog apolipoprotein(a).";
RL   J. Biol. Chem. 270:24004-24009(1995).
RN   [2]
RP   REVISIONS.
RA   Lawn R.M.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U33171; AAC48717.1; -.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL        1     19       BY SIMILARITY.
FT   CHAIN        20    810       PLASMINOGEN.
FT   CHAIN        20    582       PLASMIN HEAVY CHAIN A (BY SIMILARITY).
FT   CHAIN       583    810       PLASMIN LIGHT CHAIN B (BY SIMILARITY).
FT   DOMAIN      583    810       SERINE PROTEASE.
FT   DOMAIN      103    181       KRINGLE 1.
FT   DOMAIN      185    262       KRINGLE 2.
FT   DOMAIN      275    352       KRINGLE 3.
```

```
FT   DOMAIN      379    456       KRINGLE 4.
FT   DOMAIN      482    561       KRINGLE 5.
FT   ACT_SITE    622    622       CHARGE RELAY SYSTEM.
FT   ACT_SITE    665    665       CHARGE RELAY SYSTEM.
FT   ACT_SITE    760    760       CHARGE RELAY SYSTEM.
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   810 AA;  90902 MW;  8E75780946017A16 CRC64;

  Query Match             87.0%;  Score 120;  DB 1;  Length 810;
  Best Local Similarity   82.6%;  Pred. No. 1.1e-09;
  Matches   19;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||| ||| ||||||||:||
Db     533 RNPDGDANGPWCYTTNPRKLFDY 555


RESULT 6
APOA_HUMAN
ID   APOA_HUMAN     STANDARD;      PRT;  4548 AA.
AC   P08519;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN   LPA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88039109; PubMed=3670400;
RA   McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA   Fless G.M., Scanu A.M., Lawn R.M.;
RT   "cDNA sequence of human apolipoprotein(a) is homologous to
RT   plasminogen.";
RL   Nature 330:132-137(1987).
RN   [2]
RP   SERINE PROTEASE ACTIVITY.
RX   MEDLINE=90076123; PubMed=2531657;
RA   Salonen E.-M., Jauhiainen M., Zardi L., Vaheri A., Ehnholm C.;
RT   "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT   activity capable of cleaving it.";
RL   EMBO J. 8:4035-4040(1989).
RN   [3]
RP   REVIEW.
RX   MEDLINE=90049223; PubMed=2530631;
RA   Utermann G.;
RT   "The mysteries of lipoprotein(a).";
RL   Science 246:904-910(1989).
RN   [4]
RP   CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX   MEDLINE=21303595; PubMed=11294842;
RA   Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT   "Structural elucidation of the N- and O-glycans of human
RT   apolipoprotein(a): role of o-glycans in conferring protease
RT   resistance.";
RL   J. Biol. Chem. 276:22200-22208(2001).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX   MEDLINE=96217891; PubMed=8642595;
RA   Mikol V., Lograsso P.V., Boettcher B.R.;
RT   "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT   complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RT   acid: existence of novel and expected binding modes.";
RL   J. Mol. Biol. 256:751-761(1996).
RN   [6]
RP   VARIANT ARG-4193.
RX   MEDLINE=95002201; PubMed=7918682;
RA   Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT   "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT   associated with a lysine binding defect in Lp(a).";
RL   Biochim. Biophys. Acta 1227:41-45(1994).
CC   -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC       (Lp(a)). It has serine proteinase activity and is able of
CC       autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC       Lp(a) may be a ligand for megalin/Gp 330.
CC   -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC       decorin.
CC   -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC       structures present in either a mono- or disialylated state. The
CC       O-glycans are mostly (80%) represented by the monosialylated core
CC       type I structure, NeuNAcalpha2-3Galbeta1-3GalNAc, with smaller
CC       amounts of disialylated and non-sialylated O-glycans also
CC       detected.
CC   -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC       naturally occuring proteolytic fragments is correlated with
CC       atherosclerosis. Homology with plasminogen kringles IV and V is
CC       thought to underlie the atherogenicity of the protein, because the
CC       fragments are competing with plasminogen for fibrin(ogen) binding.
CC   -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC       leading to the formation of the so called mini-Lp(a). Apo(a)
CC       fragments accumulate in atherosclerotic lesions, where they may
CC       promote thrombogenesis. O-glycosylation may limit the extent of
CC       proteolytic fragmentation.
CC   -!- SIMILARITY: CONTAINS 38 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06290; CAA29618.1; -.
DR   PIR; S00657; S00657.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.226; -.
DR   MIM; 152200; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 38.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 38.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 38.
DR   PROSITE; PS50070; KRINGLE_2; 38.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW   Kringle; Repeat; Atherosclerosis; Signal; Polymorphism.
FT   SIGNAL        1     19
FT   CHAIN        20   4548       APOLIPOPROTEIN(A).
FT   DOMAIN       20    130       KRINGLE TYPE IV, 1.
FT   DOMAIN      131    244       KRINGLE TYPE IV, 2.
FT   DOMAIN      245    358       KRINGLE TYPE IV, 3.
FT   DOMAIN      359    472       KRINGLE TYPE IV, 4.
FT   DOMAIN      473    586       KRINGLE TYPE IV, 5.
FT   DOMAIN      587    700       KRINGLE TYPE IV, 6.
FT   DOMAIN      701    814       KRINGLE TYPE IV, 7.
FT   DOMAIN      815    928       KRINGLE TYPE IV, 8.
FT   DOMAIN      929   1042       KRINGLE TYPE IV, 9.
FT   DOMAIN     1043   1156       KRINGLE TYPE IV, 10.
FT   DOMAIN     1157   1270       KRINGLE TYPE IV, 11.
FT   DOMAIN     1271   1384       KRINGLE TYPE IV, 12.
FT   DOMAIN     1385   1498       KRINGLE TYPE IV, 13.
FT   DOMAIN     1499   1612       KRINGLE TYPE IV, 14.
FT   DOMAIN     1613   1726       KRINGLE TYPE IV, 15.
```

```
FT   DOMAIN      1727   1840       KRINGLE TYPE IV, 16.
FT   DOMAIN      1841   1954       KRINGLE TYPE IV, 17.
FT   DOMAIN      1955   2068       KRINGLE TYPE IV, 18.
FT   DOMAIN      2069   2182       KRINGLE TYPE IV, 19.
FT   DOMAIN      2183   2296       KRINGLE TYPE IV, 20.
FT   DOMAIN      2297   2410       KRINGLE TYPE IV, 21.
FT   DOMAIN      2411   2524       KRINGLE TYPE IV, 22.
FT   DOMAIN      2525   2638       KRINGLE TYPE IV, 23.
FT   DOMAIN      2639   2752       KRINGLE TYPE IV, 24.
FT   DOMAIN      2753   2866       KRINGLE TYPE IV, 25.
FT   DOMAIN      2867   2980       KRINGLE TYPE IV, 26.
FT   DOMAIN      2981   3094       KRINGLE TYPE IV, 27.
FT   DOMAIN      3095   3208       KRINGLE TYPE IV, 28.
FT   DOMAIN      3209   3322       KRINGLE TYPE IV, 29.
FT   DOMAIN      3323   3436       KRINGLE TYPE IV, 30.
FT   DOMAIN      3437   3550       KRINGLE TYPE IV, 31.
FT   DOMAIN      3551   3664       KRINGLE TYPE IV, 32.
FT   DOMAIN      3665   3770       KRINGLE TYPE IV, 33.
FT   DOMAIN      3771   3884       KRINGLE TYPE IV, 34.
FT   DOMAIN      3885   3998       KRINGLE TYPE IV, 35.
FT   DOMAIN      3999   4112       KRINGLE TYPE IV, 36.
FT   DOMAIN      4113   4226       KRINGLE TYPE IV, 37.
FT   DOMAIN      4227   4327       KRINGLE TYPE V.
FT   DOMAIN      4328   4548       SERINE PROTEASE.
FT   ACT_SITE    4369   4369       CHARGE RELAY SYSTEM.
FT   ACT_SITE    4412   4412       CHARGE RELAY SYSTEM.
FT   ACT_SITE    4498   4498       CHARGE RELAY SYSTEM.
FT   VARIANT     4193   4193       W -> R (LOSS OF LYSINE-SEPHAROSE
FT                                 BINDING).
FT                                 /FTId=VAR_006633.
SQ   SEQUENCE   4548 AA;  501313 MW;  96921BE96A465C5F CRC64;

  Query Match             84.8%;  Score 117;  DB 1;  Length 4548;
  Best Local Similarity   78.3%;  Pred. No. 1.8e-08;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||||: ||| || |||||:||
Db    4279 RNPDGDINGPWCYTMNPRKLFDY 4301


RESULT 7
PLMN_HORSE
ID   PLMN_HORSE     STANDARD;      PRT;   338 AA.
AC   P80010;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=92052077; PubMed=1946332;
RA   Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT   "Complete amino acid sequence of equine miniplasminogen.";
RL   Protein Seq. Data Anal. 4:69-74(1991).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
DR   PIR; S17527; S17527.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER       1      1
FT   CHAIN        <1    108       PLASMIN HEAVY CHAIN A.
FT   CHAIN       109    338       PLASMIN LIGHT CHAIN B.
FT   DOMAIN        9     88       KRINGLE 5.
FT   DOMAIN      109    338       SERINE PROTEASE.
FT   DISULFID      9     88       BY SIMILARITY.
FT   DISULFID     30     71       BY SIMILARITY.
FT   DISULFID     59     83       BY SIMILARITY.
FT   DISULFID     95    213       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    105    113       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    135    151       BY SIMILARITY.
FT   DISULFID    227    294       BY SIMILARITY.
FT   DISULFID    257    273       BY SIMILARITY.
FT   DISULFID    284    312       BY SIMILARITY.
FT   ACT_SITE    150    150       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    193    193       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    288    288       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE        157    157       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        191    191       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        269    269       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE        282    282       SITE OF SUBSTRATE SPECIFICITY
FT                                (BY SIMILARITY).
SQ   SEQUENCE   338 AA;  37132 MW;  8E9E56B5C5CDBE01 CRC64;

  Query Match             82.6%;  Score 114;  DB 1;  Length 338;
  Best Local Similarity   78.3%;  Pred. No. 3.3e-09;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| || ||:||:||
Db      60 RNPDGDVNGPWCYTMNPQKLFDY 82


RESULT 8
PLMN_PIG
ID   PLMN_PIG       STANDARD;      PRT;   790 AA.
AC   P06867;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE OF 1-560.
RA   Schaller J., Marti T., Roesselet S.J., Kaempfer U., Rickli E.E.;
RT   "Amino acid sequence of the heavy chain of porcine plasmin. Comparison
RT   of the carbohydrate attachment sites with the human and bovine
RT   species.";
RL   Fibrinolysis 1:91-102(1987).
RN   [2]
```

```
RP   SEQUENCE OF 450-790.
RX   MEDLINE=85203907; PubMed=3846533;
RA   Marti T., Schaller J., Rickli E.E.;
RT   "Determination of the complete amino-acid sequence of porcine
RT   miniplasminogen.";
RL   Eur. J. Biochem. 149:279-285(1985).
RN   [3]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=88185329; PubMed=3356193;
RA   Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA   Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT   "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT   plasminogen. Species specificity in relation to sialylation and
RT   fucosylation patterns.";
RL   Eur. J. Biochem. 173:57-63(1988).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLLACTOSAMINE, SIALIC ACID AND
CC       IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAC
CC       DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
CC       (MICROHETEROGENEITY).
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
DR   PIR; A25834; A25834.
DR   PIR; S03733; S03733.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   GlycoSuiteDB; P06867; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   CHAIN         1    560       PLASMIN HEAVY CHAIN A.
FT   CHAIN       561    790       PLASMIN LIGHT CHAIN B.
FT   DOMAIN      561    790       SERINE PROTEASE.
FT   DOMAIN       84    162       KRINGLE 1.
FT   DOMAIN      166    243       KRINGLE 2.
FT   DOMAIN      256    333       KRINGLE 3.
FT   DOMAIN      358    435       KRINGLE 4.
FT   DOMAIN      461    540       KRINGLE 5.
FT   ACT_SITE    602    602       CHARGE RELAY SYSTEM.
FT   ACT_SITE    645    645       CHARGE RELAY SYSTEM.
FT   ACT_SITE    740    740       CHARGE RELAY SYSTEM.
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .).
FT                                /FTId=CAR_000019.
FT   CARBOHYD    340    340       O-LINKED (GALNAC. . .).
FT                                /FTId=CAR_000020.
SQ   SEQUENCE   790 AA;  88592 MW;  F04EA06E74BCD58E CRC64;

  Query Match             81.9%;  Score 113;  DB 1;  Length 790;
  Best Local Similarity   78.3%;  Pred. No. 1.1e-08;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||  ||| |||||:||:|| 
Db     512 RNPDGDDNGPWCYTTNPQKLFDY 534


RESULT 9
PLMN_BOVIN
ID   PLMN_BOVIN     STANDARD;      PRT;   812 AA.
AC   P06868; Q28162;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen precursor (EC 3.4.21.7).
GN   PLG.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Berglund L., Andersen M.D., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen cDNA.";
RL   Int. Dairy J. 5:593-603(1995).
RN   [2]
RP   SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=85203906; PubMed=3846532;
RA   Schaller J., Moser P.W., Dannegger-Muller G.A.K., Rosselet S.J.,
RA   Kampfer U., Rickli E.E.;
RT   "Complete amino acid sequence of bovine plasminogen. Comparison with
RT   human plasminogen.";
RL   Eur. J. Biochem. 149:267-278(1985).
RN   [3]
RP   SEQUENCE OF 706-812 FROM N.A.
RX   MEDLINE=85023311; PubMed=6148961;
RA   Malinowski D.P., Sadler J.E., Davie E.W.;
RT   "Characterization of a complementary deoxyribonucleic acid coding for
RT   human and bovine plasminogen.";
RL   Biochemistry 23:4243-4250(1984).
RN   [4]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=88185329; PubMed=3356193;
RA   Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA   Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT   "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT   plasminogen. Species specificity in relation to sialylation and
RT   fucosylation patterns.";
RL   Eur. J. Biochem. 173:57-63(1988).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLLACTOSAMINE AND SIALIC ACID.
CC       O-LINKED GLYCANS CONSIST OF GAL-GALNAC DISACCHARIDE WITH IS
CC       MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
```

```
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79402; CAA55939.1; -.
DR   EMBL; K02935; AAA30714.1; -.
DR   PIR; A25835; PLBO.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.233; -.
DR   GlycoSuiteDB; P06868; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 5.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 5.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 5.
DR   PROSITE; PS50070; KRINGLE_2; 5.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen; Signal.
FT   SIGNAL       1     26
FT   CHAIN       27    812       PLASMINOGEN.
FT   CHAIN       27    583       PLASMIN HEAVY CHAIN A.
FT   CHAIN      584    812       PLASMIN LIGHT CHAIN B.
FT   DOMAIN     110    188       KRINGLE 1.
FT   DOMAIN     192    269       KRINGLE 2.
FT   DOMAIN     282    359       KRINGLE 3.
FT   DOMAIN     384    461       KRINGLE 4.
FT   DOMAIN     485    564       KRINGLE 5.
FT   DOMAIN     584    812       SERINE PROTEASE.
FT   CARBOHYD   315    315       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000014.
FT   CARBOHYD   365    365       O-LINKED (GALNAC. . .).
FT                               /FTId=CAR_000015.
FT   ACT_SITE   624    624       CHARGE RELAY SYSTEM.
FT   ACT_SITE   667    667       CHARGE RELAY SYSTEM.
FT   ACT_SITE   762    762       CHARGE RELAY SYSTEM.
FT   CONFLICT   335    335       N -> D (IN REF. 2).
FT   CONFLICT   516    516       Q -> H (IN REF. 2).
FT   CONFLICT   555    555       P -> L (IN REF. 2).
FT   CONFLICT   744    744       T -> R (IN REF. 3).
SQ   SEQUENCE   812 AA;  91216 MW;  38A6AA691E220946 CRC64;

  Query Match             80.4%;  Score 111;  DB 1;  Length 812;
  Best Local Similarity   78.3%;  Pred. No. 2.2e-08;
  Matches   18;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           ||||||| ||| || |||| :||
Db     536 RNPDGDVNGPWCYTMNPRKPFDY 558


RESULT 10
PLMN_CANFA
ID   PLMN_CANFA     STANDARD;      PRT;   333 AA.
AC   P80009;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=90175323; PubMed=2626424;
RA   Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT   "Complete amino acid sequence of canine miniplasminogen.";
RL   Protein Seq. Data Anal. 2:445-450(1989).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF
CC       STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
DR   HSSP; P00747; 5HPG.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_TER      1      1
FT   CHAIN       <1    103       PLASMIN HEAVY CHAIN A.
FT   CHAIN      104    333       PLASMIN LIGHT CHAIN B.
FT   DOMAIN       4     83       KRINGLE 5.
FT   DOMAIN     104    333       SERINE PROTEASE.
FT   DISULFID     4     83       BY SIMILARITY.
FT   DISULFID    25     66       BY SIMILARITY.
FT   DISULFID    54     78       BY SIMILARITY.
FT   DISULFID    90    208       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   100    108       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   130    146       BY SIMILARITY.
FT   DISULFID   222    289       BY SIMILARITY.
FT   DISULFID   252    268       BY SIMILARITY.
FT   DISULFID   279    307       BY SIMILARITY.
FT   ACT_SITE   145    145       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   188    188       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   283    283       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE       152    152       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE       186    186       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE       264    264       STREPTOKINASE-BINDING SITE (PROBABLE).
FT   SITE       277    277       SITE OF SUBSTRATE SPECIFICITY
FT                               (BY SIMILARITY).
SQ   SEQUENCE   333 AA;  36678 MW;  C8C0271B6C6AC8D4 CRC64;
```

```
  Query Match             79.7%;  Score 110;  DB 1;  Length 333;
  Best Local Similarity   78.3%;  Pred. No. 1.2e-08;
  Matches   18;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||| ||| || | |||:||
Db     55 RNPDGDVNGPWCYTMNQRKLFDY 77


RESULT 11
HGFL_HUMAN
ID   HGFL_HUMAN     STANDARD;      PRT;   711 AA.
AC   P26927; Q13350; Q14870;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor-like protein precursor (Macrophage
DE   stimulatory protein) (MSP) (Macrophage stimulating protein).
GN   MST1 OR HGFL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92002016; PubMed=1655021;
RA   Han S., Stuart L.A., Friezner Degen S.J.;
RT   "Characterization of the DNF15S2 locus on human chromosome 3:
RT   identification of a gene coding for four kringle domains with
RT   homology to hepatocyte growth factor.";
RL   Biochemistry 30:9768-9780(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93340141; PubMed=8393443;
RA   Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT   "Cloning, sequencing, and expression of human macrophage stimulating
RT   protein (MSP, MST1) confirms MSP as a member of the family of kringle
RT   proteins and locates the MSP gene on chromosome 3.";
RL   J. Biol. Chem. 268:15461-15468(1993).
CC   -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC       CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC       CONSERVED.
CC   -!- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE
CC       HELD TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74178; AAA50165.1; -.
DR   EMBL; U37055; AAC50471.1; -.
DR   EMBL; L11924; AAA59872.1; -.
DR   PIR; A40331; A40331.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.975; -.
DR   MIM; 142408; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Kringle; Glycoprotein; Serine protease homolog; Signal;
KW   Polymorphism.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    711       HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
FT   DOMAIN       32    109       PAP.
FT   DOMAIN      110    186       KRINGLE 1.
FT   DOMAIN      191    268       KRINGLE 2.
FT   DOMAIN      283    361       KRINGLE 3.
FT   DOMAIN      370    448       KRINGLE 4.
FT   DOMAIN      484    711       SERINE PROTEASE-LIKE.
FT   DISULFID     56     78       BY SIMILARITY.
FT   DISULFID     60     66       BY SIMILARITY.
FT   DISULFID    110    186       BY SIMILARITY.
FT   DISULFID    131    169       BY SIMILARITY.
FT   DISULFID    157    181       BY SIMILARITY.
FT   DISULFID    191    268       BY SIMILARITY.
FT   DISULFID    194    324       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    212    251       BY SIMILARITY.
FT   DISULFID    240    263       BY SIMILARITY.
FT   DISULFID    283    361       BY SIMILARITY.
FT   DISULFID    304    343       BY SIMILARITY.
FT   DISULFID    332    355       BY SIMILARITY.
FT   DISULFID    370    448       BY SIMILARITY.
FT   DISULFID    391    431       BY SIMILARITY.
FT   DISULFID    419    443       BY SIMILARITY.
FT   DISULFID    468    588       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    507    523       BY SIMILARITY.
FT   DISULFID    602    667       BY SIMILARITY.
FT   DISULFID    632    646       BY SIMILARITY.
FT   DISULFID    657    685       BY SIMILARITY.
FT   CARBOHYD     72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      13     13       Y -> C.
FT                                /FTId=VAR_006631.
FT   VARIANT     212    212       C -> F.
FT                                /FTId=VAR_006632.
FT   CONFLICT    623    623       L -> F (IN REF. 2).
SQ   SEQUENCE   711 AA;  80379 MW;  596ED21F180290E4 CRC64;

  Query Match             66.7%;  Score 92;  DB 1;  Length 711;
  Best Local Similarity   65.2%;  Pred. No. 1.1e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 RNPDGDVGGPWAYTTNPRKLYDY 23
          ||||||  ||| || :||  :||
Db    420 RNPDGDSHGPWCYTMDPRTPFDY 442


RESULT 12
HGFL_MOUSE
ID   HGFL_MOUSE     STANDARD;      PRT;   716 AA.
AC   P26928;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor-like protein precursor (Macrophage
DE   stimulatory protein) (MSP).
GN   MST1 OR HGFL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=BALB/C; TISSUE=Liver;
RX   MEDLINE=92002017; PubMed=1832957;
RA   Friezner Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT   "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT   growth factor-like protein: expression during development.";
RL   Biochemistry 30:9781-9791(1991).
CC   -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC       CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC       CONSERVED.
CC   -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC       ADRENAL.
CC   -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC       JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC       STABLE AFTERWARDS.
CC   -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC       HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC       POLYPEPTIDES.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74180; AAA50166.1; -.
DR   EMBL; M74181; AAA50167.1; -.
DR   HSSP; P00747; 1KRN.
DR   MEROPS; S01.975; -.
DR   MGD; MGI:96080; Hgfl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Kringle; Glycoprotein; Serine protease homolog; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    716       HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
FT   DOMAIN      19    109       PAP.
FT   DOMAIN     110    186       KRINGLE 1.
FT   DOMAIN     191    268       KRINGLE 2.
FT   DOMAIN     292    370       KRINGLE 3.
FT   DOMAIN     379    457       KRINGLE 4.
FT   DOMAIN     489    716       SERINE PROTEASE-LIKE.
FT   DISULFID    56     78       BY SIMILARITY.
FT   DISULFID    60     66       BY SIMILARITY.
FT   DISULFID   110    186       BY SIMILARITY.
FT   DISULFID   131    169       BY SIMILARITY.
FT   DISULFID   157    181       BY SIMILARITY.
FT   DISULFID   191    268       BY SIMILARITY.
FT   DISULFID   194    333       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   212    251       BY SIMILARITY.
FT   DISULFID   240    263       BY SIMILARITY.
FT   DISULFID   292    370       BY SIMILARITY.
FT   DISULFID   313    352       BY SIMILARITY.
FT   DISULFID   341    364       BY SIMILARITY.
FT   DISULFID   379    457       BY SIMILARITY.
FT   DISULFID   400    440       BY SIMILARITY.
FT   DISULFID   428    452       BY SIMILARITY.
FT   DISULFID   477    593       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   512    528       BY SIMILARITY.
FT   DISULFID   607    672       BY SIMILARITY.
FT   DISULFID   637    651       BY SIMILARITY.
FT   DISULFID   662    690       BY SIMILARITY.
FT   CARBOHYD    72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   305    305       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   620    620       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    19     19       P -> Q (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   716 AA;  80588 MW;  BBCE02EF85213ACC CRC64;

  Query Match             65.2%;  Score 90;  DB 1;  Length 716;
  Best Local Similarity   65.2%;  Pred. No. 2.1e-05;
  Matches   15;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||||  ||| || :|  |:||
Db     429 RNPDGDSHGPWCYTLDPDILFDY 451


RESULT 13
HGF_RAT
ID   HGF_RAT        STANDARD;      PRT;   728 AA.
AC   P17945;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hepatocyte growth factor precursor (Scatter factor) (SF)
DE   (Hepatopoeitin A).
GN   HGF.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=90222197; PubMed=2139229;
RA   Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
RA   Shimizu S., Nakamura T.;
RT   "Deduced primary structure of rat hepatocyte growth factor and
RT   expression of the mRNA in rat tissues.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=91031482; PubMed=2146117;
RA   Okajima A., Miyazawa K., Kitamura N.;
RT   "Primary structure of rat hepatocyte growth factor and induction of
RT   its mRNA during liver regeneration following hepatic injury.";
RL   Eur. J. Biochem. 193:375-381(1990).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC       AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC       IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90102; BAA14133.1; -.
DR   EMBL; X54400; CAA38266.1; -.
```

```
DR   PIR; S13211; S13211.
DR   PIR; A35644; A35644.
DR   HSSP; P14210; 1BHT.
DR   MEROPS; S01.978; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW   Signal.
FT   SIGNAL       1     32       BY SIMILARITY.
FT   CHAIN       33    495       HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT   CHAIN      496    728       HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT   MOD_RES      33     33       PYRROLIDONE CARBOXYLIC ACID
FT                                (BY SIMILARITY).
FT   DOMAIN       33    128       PAP.
FT   DOMAIN      129    207       KRINGLE 1.
FT   DOMAIN      212    289       KRINGLE 2.
FT   DOMAIN      306    384       KRINGLE 3.
FT   DOMAIN      392    470       KRINGLE 4.
FT   DOMAIN      496    728       SERINE PROTEASE-LIKE.
FT   DISULFID     71     97       BY SIMILARITY.
FT   DISULFID     75     85       BY SIMILARITY.
FT   DISULFID    488    607       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    569    569       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    656    656       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   728 AA;  82905 MW;  3E0BF1F96ADCEDFF CRC64;

  Query Match             59.4%;  Score 82;  DB 1;  Length 728;
  Best Local Similarity   60.9%;  Pred. No. 0.0003;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || ||   :||
Db     442 RNPDDDAHGPWCYTGNPLVPWDY 464


RESULT 14
HGF_HUMAN
ID   HGF_HUMAN      STANDARD;      PRT;   728 AA.
AC   P14210; Q9UDU6; Q9BYL9;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor precursor (Scatter factor) (SF)
DE   (Hepatopoeitin A).
GN   HGF OR HPTA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91340155; PubMed=1831432;
RA   Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT   "Organization of the human hepatocyte growth factor-encoding gene.";
RL   Gene 102:213-219(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=89392017; PubMed=2528952;
RA   Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA   Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RA   Gohda E., Daikuhara Y., Kitamura N.;
RT   "Molecular cloning and sequence analysis of cDNA for human hepatocyte
RT   growth factor.";
RL   Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=91025062; PubMed=2145836;
RA   Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA   Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT   "Isolation and expression of cDNA for different forms of hepatocyte
RT   growth factor from human leukocyte.";
RL   Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN   [4]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RC   TISSUE=Liver;
RX   MEDLINE=90066676; PubMed=2531289;
RA   Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA   Sugimura A., Tashiro K., Shimizu S.;
RT   "Molecular cloning and expression of human hepatocyte growth factor.";
RL   Nature 342:440-443(1989).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryonic fibroblast;
RX   MEDLINE=91334393; PubMed=1831266;
RA   Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA   Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA   Birchmeier W.;
RT   "Evidence for the identity of human scatter factor and human
RT   hepatocyte growth factor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Courtney L., Elliot G., Angell S.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 249-695 FROM N.A.
RX   MEDLINE=91369928; PubMed=1832556;
RA   Miyazawa K., Kitamura A., Kitamura N.;
RT   "Structural organization and the transcription initiation site of the
RT   human hepatocyte growth factor gene.";
RL   Biochemistry 30:9170-9176(1991).
RN   [8]
RP   SIGNAL SEQUENCE CLEAVAGE SITE.
RX   MEDLINE=91207365; PubMed=1826837;
RA   Yoshiyama Y., Arakakai N., Naka D., Takahashi K., Hirono S., Kondo J.,
RA   Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
RA   Hishida T., Daikuhara Y.;
RT   "Identification of the N-terminal residue of the heavy chain of both
RT   native and recombinant human hepatocyte growth factor.";
RL   Biochem. Biophys. Res. Commun. 175:660-667(1991).
RN   [9]
RP   CARBOHYDRATE-LINKAGE SITE 476.
RX   MEDLINE=93129192; PubMed=1482348;
RA   Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
RA   Nakamura T., Shimizu S.;
RT   "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
RT   on the alpha chain.";
RL   Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
RN   [10]
RP   MUTAGENESIS.
RX   MEDLINE=92331602; PubMed=1321034;
RA   Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
RA   Baker J.B., Godowski P.J.;
RT   "Structure-function analysis of hepatocyte growth factor:
RT   identification of variants that lack mitogenic activity yet retain
RT   high affinity receptor binding.";
RL   EMBO J. 11:2503-2510(1992).
RN   [11]
```

```
RP   STRUCTURE BY NMR OF 31-127.
RX   MEDLINE=98154323; PubMed=9493272;
RA   Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
RA   Rubin J.S., Bottaro D.P., Byrd R.A.;
RT   "The solution structure of the N-terminal domain of hepatocyte growth
RT   factor reveals a potential heparin-binding site.";
RL   Structure 6:109-116(1998).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
RX   MEDLINE=99036858; PubMed=9817840;
RA   Ultsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
RT   "Crystal structure of the NK1 fragment of human hepatocyte growth
RT   factor at 2.0-A resolution.";
RL   Structure 6:1383-1393(1998).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC       AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC       IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D90334; BAA14348.1; -.
DR   EMBL; D90318; BAA14348.1; JOINED.
DR   EMBL; D90319; BAA14348.1; JOINED.
DR   EMBL; D90320; BAA14348.1; JOINED.
DR   EMBL; D90322; BAA14348.1; JOINED.
DR   EMBL; D90323; BAA14348.1; JOINED.
DR   EMBL; D90324; BAA14348.1; JOINED.
DR   EMBL; D90325; BAA14348.1; JOINED.
DR   EMBL; D90326; BAA14348.1; JOINED.
DR   EMBL; D90327; BAA14348.1; JOINED.
DR   EMBL; D90328; BAA14348.1; JOINED.
DR   EMBL; D90329; BAA14348.1; JOINED.
DR   EMBL; D90330; BAA14348.1; JOINED.
DR   EMBL; D90331; BAA14348.1; JOINED.
DR   EMBL; D90332; BAA14348.1; JOINED.
DR   EMBL; D90333; BAA14348.1; JOINED.
DR   EMBL; M29145; AAA52650.1; -.
DR   EMBL; M60718; AAA52648.1; -.
DR   EMBL; X16323; CAA34387.1; -.
DR   EMBL; M73239; AAA64239.1; -.
DR   EMBL; M73240; AAA64297.1; -.
DR   EMBL; AC004960; AAC71655.1; -.
DR   EMBL; M75983; AAG53460.1; -.
DR   EMBL; M75972; AAG53460.1; JOINED.
DR   EMBL; M75973; AAG53460.1; JOINED.
DR   EMBL; M75974; AAG53460.1; JOINED.
DR   EMBL; M75975; AAG53460.1; JOINED.
DR   EMBL; M75976; AAG53460.1; JOINED.
DR   EMBL; M75977; AAG53460.1; JOINED.
DR   EMBL; M75978; AAG53460.1; JOINED.
DR   EMBL; M75979; AAG53460.1; JOINED.
DR   EMBL; M75980; AAG53460.1; JOINED.
DR   EMBL; M75981; AAG53460.1; JOINED.
DR   EMBL; M75982; AAG53460.1; JOINED.
DR   PIR; JH0579; JH0579.
DR   PIR; S06794; S06794.
DR   PDB; 2HGF; 24-JUN-98.
DR   PDB; 1BHT; 18-NOV-98.
DR   MEROPS; S01.976; -.
DR   GlycoSuiteDB; P14210; -.
DR   MIM; 142409; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW   Signal; 3D-structure.
FT   SIGNAL       1     31
FT   CHAIN       32    494       HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT   CHAIN      495    728       HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT   MOD_RES     32     32       PYRROLIDONE CARBOXYLIC ACID.
FT   DOMAIN      32    127       PAP.
FT   DOMAIN     128    206       KRINGLE 1.
FT   DOMAIN     211    288       KRINGLE 2.
FT   DOMAIN     305    383       KRINGLE 3.
FT   DOMAIN     391    469       KRINGLE 4.
FT   DOMAIN     495    728       SERINE PROTEASE-LIKE.
FT   DISULFID    70     96
FT   DISULFID    74     84
FT   DISULFID   128    206
FT   DISULFID   149    189
FT   DISULFID   177    201
FT   DISULFID   487    604       INTERCHAIN (BY SIMILARITY).

  Query Match             58.7%;  Score 81;  DB 1;  Length 728;
  Best Local Similarity   60.9%;  Pred. No. 0.00042;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || ||   :||
Db     441 RNPDDDAHGPWCYTGNPLIPWDY 463


RESULT 15
HGF_MOUSE
ID   HGF_MOUSE      STANDARD;      PRT;   728 AA.
AC   Q08048; Q64007; Q61662;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hepatocyte growth factor precursor (Scatter factor) (SF)
DE   (Hepatopoeitin A).
GN   HGF.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 496-504.
RC   TISSUE=Mammary fibroblast;
RX   MEDLINE=94183257; PubMed=8135822;
RA   Sasaki M., Nishio M., Sasaki T., Enami J.;
RT   "Identification of mouse mammary fibroblast-derived mammary growth
RT   factor as hepatocyte growth factor.";
RL   Biochem. Biophys. Res. Commun. 199:772-779(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=94363381; PubMed=8081873;
RA   Lee C.C., Kozak C.A., Yamada K.M.;
RT   "Structure, genetic mapping, and expression of the mouse Hgf/scatter
RT   factor gene.";
```

```
RL   Cell Adhes. Commun. 1:101-111(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=94060105; PubMed=8241272;
RA   Liu Y., Michalopoulos G.K., Zarnegar R.;
RT   "Molecular cloning and characterization of cDNA encoding mouse
RT   hepatocyte growth factor.";
RL   Biochim. Biophys. Acta 1216:299-303(1993).
CC   -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC       HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC       AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC       IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC   -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY
CC       ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE
CC       LONG FORM.
CC   -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10212; BAA01064.1; -.
DR   EMBL; D10213; BAA01065.1; -.
DR   EMBL; S71816; AAB31855.1; -.
DR   EMBL; X72307; CAA51054.1; ALT_INIT.
DR   HSSP; P14210; 1BHT.
DR   MGD; MGI:96079; Hgf.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
KW   Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW   Signal; Alternative splicing.
FT   SIGNAL       1     32       BY SIMILARITY.
FT   CHAIN       33    495       HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT   CHAIN      496    728       HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT   MOD_RES      33     33       PYRROLIDONE CARBOXYLIC ACID
FT                                (BY SIMILARITY).
FT   DOMAIN       33    128       PAP.
FT   DOMAIN      129    207       KRINGLE 1.
FT   DOMAIN      212    289       KRINGLE 2.
FT   DOMAIN      306    384       KRINGLE 3.
FT   DOMAIN      392    470       KRINGLE 4.
FT   DOMAIN      496    728       SERINE PROTEASE-LIKE.
FT   DISULFID     71     97       BY SIMILARITY.
FT   DISULFID     75     85       BY SIMILARITY.
FT   DISULFID    488    607       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    569    569       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    656    656       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    163    167       MISSING (IN SHORT ISOFORM).
FT   CONFLICT    344    344       N -> K (IN REF. 2).
FT   CONFLICT    479    479       V -> L (IN REF. 2).
FT   CONFLICT    564    564       R -> H (IN REF. 3).
SQ   SEQUENCE   728 AA;  82944 MW;  A0381FC497534328 CRC64;

  Query Match             58.7%;  Score 81;  DB 1;  Length 728;
  Best Local Similarity   60.9%;  Pred. No. 0.00042;
  Matches   14;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| || ||   :||
Db     442 RNPDDDAHGPWCYTGNPLIPWDY 464


RESULT 16
THRB_BOVIN
ID   THRB_BOVIN     STANDARD;      PRT;   625 AA.
AC   P00735;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88245190; PubMed=3379642;
RA   Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
RT   "Structure and evolution of the bovine prothrombin gene.";
RL   J. Mol. Biol. 200:31-45(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84203525; PubMed=6326805;
RA   McGillivray R.T.A., Davie E.W.;
RT   "Characterization of bovine prothrombin mRNA and its translation
RT   product.";
RL   Biochemistry 23:1626-1634(1984).
RN   [3]
RP   SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RA   Magnusson S., Sottrup-Jensen L., Petersen T.E., Claeys H.;
RL   (In) Hemker H.C., Veltkamp J.J. (eds.);
RL   Boerhaave symposium on prothrombin and related coagulation factors,
RL   pp.25-46, Leiden University Press, Leiden (1975).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=86296631; PubMed=3741841;
RA   Park C.H., Tulinsky A.;
RT   "Three-dimensional structure of the kringle sequence: structure of
RT   prothrombin fragment 1.";
RL   Biochemistry 25:3977-3982(1986).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=91311686; PubMed=1856869;
RA   Seshadri T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;
RT   "Structure of bovine prothrombin fragment 1 refined at 2.25-A
RT   resolution.";
RL   J. Mol. Biol. 220:481-494(1991).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX   MEDLINE=92190185; PubMed=1547238;
RA   Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;
RT   "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-
RT   prothrombin fragment 1.";
RL   Biochemistry 31:2554-2566(1992).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=92218459; PubMed=1560020;
RA   Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
RT   "The structure of residues 7-16 of the A alpha-chain of human
```

```
RT   fibrinogen bound to bovine thrombin at 2.3-A resolution.";
RL   J. Biol. Chem. 267:7911-7920(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=92389319; PubMed=1518046;
RA   Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
RA   Martin P.D., Edwards B.F.P., Bode W.;
RT   "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes
RT   formed with the benzamidine and arginine-based thrombin inhibitors
RT   NAPAP, 4-TAPAP and MQPA. A starting point for improving
RT   antithrombotics.";
RL   J. Mol. Biol. 226:1085-1089(1992).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
RX   MEDLINE=97102783; PubMed=8947023;
RA   van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
RA   Hoffken W., Huber R.;
RT   "The ornithodorin-thrombin crystal structure, a key to the TAP
RT   enigma?";
RL   EMBO J. 15:6011-6017(1996).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
RX   MEDLINE=98004486; PubMed=9342325;
RA   Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
RA   Huber R., Bode W.;
RT   "Structure of the thrombin complex with triabin, a lipocalin-like
RT   exosite-binding inhibitor derived from a triatomine bug.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
RN   [11]
RP   GENE STRUCTURE.
RX   MEDLINE=86077733; PubMed=3000440;
RA   Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;
RT   "Characterization of the bovine prothrombin gene.";
RL   Biochemistry 24:6854-6861(1985).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   -!- DATABASE: NAME=ProZyme technical fact sheet;
CC       WWW="http://www.prozyme.com/technical/thrombindata.html".
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V00135; CAA23451.1; -.
DR   EMBL; J00041; AAA30781.1; -.
DR   PIR; A00915; TBBO.
DR   PIR; S02537; S02537.
DR   PDB; 1BBR; 31-JAN-94.
DR   PDB; 1ETR; 31-JAN-94.
DR   PDB; 1ETS; 31-JAN-94.
DR   PDB; 1ETT; 31-JAN-94.
DR   PDB; 1HRT; 31-JAN-94.
DR   PDB; 2PF1; 31-JAN-94.
DR   PDB; 2PF2; 31-JAN-94.
DR   PDB; 2SPT; 31-MAY-94.
DR   PDB; 1MKW; 07-JUL-97.
DR   PDB; 1MKX; 07-JUL-97.
DR   PDB; 1TBQ; 14-OCT-96.
DR   PDB; 1TBR; 14-OCT-96.
DR   PDB; 1TOC; 23-JUL-97.
DR   PDB; 1VIT; 21-APR-97.
DR   PDB; 1YCP; 06-MAY-98.
DR   PDB; 1A0H; 17-JUN-98.
DR   PDB; 1AVG; 16-FEB-99.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal; 3D-structure.
FT   SIGNAL       1    24       POTENTIAL.
FT   PROPEP      25    43
FT   CHAIN       44   625       PROTHROMBIN.
FT   PEPTIDE     44   199       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE    200   317       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN      318   366       THROMBIN LIGHT CHAIN (A).
FT   CHAIN      367   625       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN     109   187       KRINGLE 1.
FT   DOMAIN     214   292       KRINGLE 2.
FT   DOMAIN     367   625       SERINE PROTEASE.
FT   SITE       199   200       CLEAVAGE (BY THROMBIN).
FT   SITE       317   318       CLEAVAGE (BY FACTOR XA).
FT   SITE       366   367       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE   409   409       CHARGE RELAY SYSTEM.
FT   ACT_SITE   465   465       CHARGE RELAY SYSTEM.
FT   ACT_SITE   571   571       CHARGE RELAY SYSTEM.
FT   MOD_RES     50    50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     51    51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     58    58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60    60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     63    63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     64    64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     69    69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     70    70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     73    73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     76    76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD   120   120       N-LINKED (GLCNAC. . .).

  Query Match           58.0%;  Score 80;  DB 1;  Length 625;
  Best Local Similarity 70.6%;  Pred. No. 0.0005;
```

```
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           |||||  : ||| |||:|
Db     159 RNPDGSITGPWCYTTSP 175


RESULT 17
PLMN_RAT
ID   PLMN_RAT       STANDARD;      PRT;   169 AA.
AC   Q01177;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragment).
GN   PLG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=91250378; PubMed=1645711;
RA   Kanalas J.J., Makker S.P.;
RT   "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT   receptor site for plasminogen.";
RL   J. Biol. Chem. 266:10825-10829(1991).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC       GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC       ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC       AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC       LAMININ AND VON WILLEBRAND FACTOR.
CC   -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC       ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC       FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC   -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC       IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC   -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62832; AAA41884.1; -.
DR   PIR; A40522; A40522.
DR   HSSP; P00747; 1PMK.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; PARTIAL.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle.
FT   NON_TER       1      1
FT   DOMAIN       <1     10       KRINGLE 3 (BY SIMILARITY).
FT   DOMAIN       34    112       KRINGLE 4 (BY SIMILARITY).
FT   DOMAIN      139   >169       KRINGLE 5 (BY SIMILARITY).
FT   DISULFID     34    112       BY SIMILARITY.
FT   DISULFID     55     95       BY SIMILARITY.
FT   DISULFID     83    107       BY SIMILARITY.
FT   NON_TER     169    169
SQ   SEQUENCE   169 AA;  18401 MW;  77A54214C49D010C CRC64;

  Query Match             55.8%;  Score 77;  DB 1;  Length 169;
  Best Local Similarity   52.2%;  Pred. No. 0.00035;
  Matches   12;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| |  ||| :||:|   ::|
Db      84 RNPDNDQRGPWCFTTDPSVRWEY 106


RESULT 18
PLMN_PETMA
ID   PLMN_PETMA     STANDARD;      PRT;   325 AA.
AC   P33574;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Plasminogen (EC 3.4.21.7) (Fragments).
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE.
RA   Affolter M., Schaller J., Rickli E.E.;
RT   "Isolation, characterization and partial amino acid sequence of
RT   lamprey plasminogen.";
RL   Protein Seq. Data Anal. 5:207-211(1993).
CC   -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC       A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC       EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC       AND INFLAMMATION.
CC   -!- SIMILARITY: CONTAINS AT LEAST 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
DR   PIR; S33879; S33879.
DR   HSSP; P00747; 1KRN.
DR   MEROPS; S01.233; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 2.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; PARTIAL.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW   Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT   NON_CONS     15     16
FT   NON_CONS     34     35
FT   NON_CONS     44     45
FT   NON_CONS     76     77
FT   NON_CONS    111    112
FT   NON_CONS    138    139
FT   NON_CONS    158    159
FT   NON_CONS    178    179
FT   NON_CONS    216    217
FT   NON_CONS    236    237
FT   NON_CONS    267    268
FT   NON_CONS    282    283
FT   NON_CONS    295    296
FT   NON_CONS    307    308
FT   NON_CONS    315    316
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  35194 MW;  1B5F0B539AC6ED3C CRC64;

  Query Match             52.9%;  Score 73;  DB 1;  Length 325;
```

```
  Best Local Similarity   70.6%;  Pred. No. 0.0026;
  Matches   12;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           |||| |  ||| || ||
Db     189 RNPDQDPQGPWCYTXNP 205


RESULT 19
THRB_RAT
ID   THRB_RAT       STANDARD;      PRT;   617 AA.
AC   P18292;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX   MEDLINE=90332426; PubMed=2377469;
RA   Dihanich M., Monard D.;
RT   "cDNA sequence of rat prothrombin.";
RL   Nucleic Acids Res. 18:4251-4251(1990).
RN   [2]
RP   SEQUENCE OF 383-617 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52835; CAA37017.1; -.
DR   EMBL; M81397; AAA42240.1; -.
DR   PIR; S10511; S10511.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   PROPEP      25     43
FT   CHAIN       44    617       PROTHROMBIN.
FT   PEPTIDE     44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE    201    323       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN      324    359       THROMBIN LIGHT CHAIN (A).
FT   CHAIN      360    617       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN     109    187       KRINGLE 1.
FT   DOMAIN     215    292       KRINGLE 2.
FT   DOMAIN     360    617       SERINE PROTEASE.
FT   SITE       200    201       CLEAVAGE (BY THROMBIN).
FT   SITE       323    324       CLEAVAGE (BY FACTOR XA).
FT   SITE       359    360       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE   402    402       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   458    458       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   564    564       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES     50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD   120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   552    552       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    61     66       BY SIMILARITY.
FT   DISULFID    91    104       BY SIMILARITY.
FT   DISULFID   109    187       BY SIMILARITY.
FT   DISULFID   130    170       BY SIMILARITY.
FT   DISULFID   158    182       BY SIMILARITY.
FT   DISULFID   215    292       BY SIMILARITY.
FT   DISULFID   236    276       BY SIMILARITY.
FT   DISULFID   264    287       BY SIMILARITY.
FT   DISULFID   332    478       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   387    403       BY SIMILARITY.
FT   DISULFID   532    546       BY SIMILARITY.
FT   DISULFID   560    590       BY SIMILARITY.
SQ   SEQUENCE   617 AA;  70411 MW;  AD27D1B71445DB1D CRC64;

  Query Match           52.9%;  Score 73;  DB 1;  Length 617;
  Best Local Similarity   64.7%;  Pred. No. 0.0051;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy       1 RNPDGDVGGPWAYTTNP 17
             ||||    ||| |||:|
Db     159 RNPDSSTSGPWCYTTDP 175


RESULT 20
THRB_MOUSE
ID   THRB_MOUSE     STANDARD;      PRT;   618 AA.
AC   P19221;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2 OR CF2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=91025551; PubMed=2222810;
RA   Friezner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA   Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse prothrombin and
RT   localization of the gene on mouse chromosome 2.";
RL   DNA Cell Biol. 9:487-498(1990).
RN   [2]
RP   SEQUENCE OF 384-618 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52308; CAA36548.1; -.
DR   EMBL; M81394; AAA40435.1; -.
DR   PIR; A35827; A35827.
DR   HSSP; P00734; 1B7X.
DR   MEROPS; S01.217; -.
DR   MGD; MGI:88380; F2.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   PROPEP       25     43
FT   CHAIN        44    618       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    324       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       325    360       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       361    618       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      361    618       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        324    325       CLEAVAGE (BY FACTOR XA).
FT   SITE        360    361       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    403    403       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    459    459       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    565    565       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    293       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    288       BY SIMILARITY.
FT   DISULFID    333    479       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    388    404       BY SIMILARITY.
FT   DISULFID    533    547       BY SIMILARITY.
FT   DISULFID    561    591       BY SIMILARITY.
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    413    413       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    553    553       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   618 AA;  70268 MW;  B89F719AAFD601E0 CRC64;

  Query Match             51.4%;  Score 71;  DB 1;  Length 618;
  Best Local Similarity   64.7%;  Pred. No. 0.0099;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy      1 RNPDGDVGGPWAYTTNP 17
          ||||    ||| |||:|
Db    159 RNPDSSTTGPWCYTTDP 175
```

```
RESULT 21
THRB_HUMAN
ID   THRB_HUMAN     STANDARD;      PRT;   622 AA.
AC   P00734;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).
GN   F2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88077877; PubMed=2825773;
RA   Degen S.J.F., Davie E.W.;
RT   "Nucleotide sequence of the gene for human prothrombin.";
RL   Biochemistry 26:6165-6177(1987).
RN   [2]
RP   SEQUENCE OF 8-622 FROM N.A.
RX   MEDLINE=83231469; PubMed=6305407;
RA   Degen S.J.F., McGillivray R.T.A., Davie E.W.;
RT   "Characterization of the complementary deoxyribonucleic acid and gene
RT   coding for human prothrombin.";
RL   Biochemistry 22:2087-2097(1983).
RN   [3]
RP   SEQUENCE OF 44-314.
RX   MEDLINE=77193964; PubMed=266717;
RA   Walz D.A., Hewett-Emmett D., Seegers W.H.;
RT   "Amino acid sequence of human prothrombin fragments 1 and 2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).
RN   [4]
RP   SEQUENCE OF 315-622.
RX   MEDLINE=77207112; PubMed=873923;
RA   Butkowski R.J., Elion J., Downing M.R., Mann K.G.;
RT   "Primary structure of human prethrombin 2 and alpha-thrombin.";
RL   J. Biol. Chem. 252:4942-4957(1977).
RN   [5]
RP   PROCESSING.
RX   MEDLINE=87008532; PubMed=3759958;
RA   Rabiet M.J., Blashill A., Furie B., Furie B.C.;
RT   "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin
RT   activation in human plasma.";
RL   J. Biol. Chem. 261:13210-13215(1986).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=90059942; PubMed=2583108;
RA   Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;
RT   "The refined 1.9 A crystal structure of human alpha-thrombin:
RT   interaction with D-Phe-Pro-Arg chloromethylketone and significance of
RT   the Tyr-Pro-Pro-Trp insertion segment.";
RL   EMBO J. 8:3467-3475(1989).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=90327074; PubMed=2374926;
RA   Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,
RA   Roitsch C., Fenton J.W. II;
RT   "The structure of a complex of recombinant hirudin and human alpha-
RT   thrombin.";
RL   Science 249:277-280(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=94350942; PubMed=8071320;
RA   Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,
RA   Correa P.E., Fenton J.W. II, Tulinsky A.;
RT   "Crystallographic structure of human gamma-thrombin.";
RL   J. Biol. Chem. 269:22000-22006(1994).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=97357286; PubMed=9214615;
RA   van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,
RA   Esmon C.T., Stubbs M.T.;
RT   "The thrombin E192Q-BPTI complex reveals gross structural
RT   rearrangements: implications for the interaction with antithrombin
RT   and thrombomodulin.";
RL   EMBO J. 16:2977-2984(1997).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.
RX   MEDLINE=99162521; PubMed=10051558;
RA   Guinto E.R., Caccia S., Rose T., Fuetterer K., Waksman G., di Cera E.;
RT   "Unexpected crucial role of residue 225 in serine proteases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).
RN   [11]
RP   VARIANT BARCELONA.
RX   MEDLINE=87033739; PubMed=3771562;
RA   Rabiet M.-J., Furie B.C., Furie B.;
RT   "Molecular defect of prothrombin Barcelona. Substitution of cysteine
RT   for arginine at residue 273.";
RL   J. Biol. Chem. 261:15045-15048(1986).
RN   [12]
RP   VARIANT FRANKFURT.
RX   MEDLINE=95313001; PubMed=7792730;
RA   Degen S.J.F., McDowell S.A., Sparks L.M., Scharrer I.;
RT   "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by
RT   substitution of Glu-466 by Ala.";
RL   Thromb. Haemost. 73:203-209(1995).
RN   [13]
RP   VARIANTS HIMI-1 AND HIMI-2.
RX   MEDLINE=93043342; PubMed=1421398;
RA   Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,
RA   Yamaguchi K.;
RT   "Prothrombin Himi: a compound heterozygote for two dysfunctional
RT   prothrombin molecules (Met-337-->Thr and Arg-388-->His).";
RL   Blood 80:2275-2280(1992).
RN   [14]
RP   VARIANT PADUA-1.
RX   MEDLINE=95169898; PubMed=7865694;
RA   James H.L., Kim D.J., Zheng D.-Q., Girolami A.;
RT   "Prothrombin Padua I: incomplete activation due to an amino acid
RT   substitution at a factor Xa cleavage site.";
RL   Blood Coagul. Fibrinolysis 5:841-844(1994).
RN   [15]
RP   VARIANT QUICK-1.
RX   MEDLINE=89207504; PubMed=3242619;
RA   Henriksen R.A., Mann K.G.;
RT   "Identification of the primary structural defect in the dysthrombin
RT   thrombin Quick I: substitution of cysteine for arginine-382.";
RL   Biochemistry 27:9160-9165(1988).
RN   [16]
RP   VARIANT QUICK-2.
RX   MEDLINE=89247398; PubMed=2719946;
RA   Henriksen R.A., Mann K.G.;
RT   "Substitution of valine for glycine-558 in the congenital dysthrombin
RT   thrombin Quick II alters primary substrate specificity.";
RL   Biochemistry 28:2078-2082(1989).
RN   [17]
RP   VARIANT SALAKTA.
RX   MEDLINE=92378975; PubMed=1354985;
RA   Miyata T., Aruga R., Umeyama H., Bezeaud A., Guillin M.-C.,
RA   Iwanaga S.;
RT   "Prothrombin Salakta: substitution of glutamic acid-466 by alanine
RT   reduces the fibrinogen clotting activity and the esterase activity.";
RL   Biochemistry 31:7457-7462(1992).
RN   [18]
RP   VARIANT TOKUSHIMA.
RX   MEDLINE=87185407; PubMed=3567158;
RA   Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,
RA   Iwanaga S.;
RT   "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan
RT   that impairs the fibrinogen clotting activity of derived thrombin
```

```
RT   Tokushima.";
RL   Biochemistry 26:1117-1122(1987).
RN   [19]
RP   VARIANT TOKUSHIMA.
RX   MEDLINE=87101511; PubMed=3801671;
RA   Inomoto T., Shirakami A., Kawauchi S., Shigekiyo T., Saito S.,
RA   Miyoshi K., Morita T., Iwanaga S.;
RT   "Prothrombin Tokushima: characterization of dysfunctional thrombin
RT   derived from a variant of human prothrombin.";
RL   Blood 69:565-569(1987).
RN   [20]
RP   VARIANT TOKUSHIMA.
RX   MEDLINE=92256895; PubMed=1349838;
RA   Iwahana H., Yoshimoto K., Shigekiyo T., Shirakami A., Saito S.,
RA   Itakura M.;
RT   "Detection of a single base substitution of the gene for prothrombin
RT   Tokushima. The application of PCR-SSCP for the genetic and molecular
RT   analysis of dysprothrombinemia.";
RL   Int. J. Hematol. 55:93-100(1992).
RN   [21]
RP   VARIANT TYPE-3.
RX   MEDLINE=83204687; PubMed=6405779;
RA   Board P.G., Shaw D.C.;
RT   "Determination of the amino acid substitution in human prothrombin
RT   type 3 (157 Glu leads to Lys) and the localization of a third
RT   thrombin cleavage site.";
RL   Br. J. Haematol. 54:245-254(1983).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- DISEASE: DEFECTS IN F2 ARE THE CAUSE OF VARIOUS FORMS OF
CC       DYSPROTHROMBINEMIA.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: IT IS NOT KNOWN WHETHER 1 OR 2 SMALLER ACTIVATION
CC       PEPTIDES, WITH ADDITIONAL CLEAVAGE AFTER 314-ARG, ARE RELEASED IN
CC       NATURAL BLOOD CLOTTING.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- MISCELLANEOUS: THE CLEAVAGE AFTER R-198, OBSERVED IN VITRO, DOES
CC       NOT OCCUR IN PLASMA.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -

  Query Match             51.4%;  Score 71;  DB 1;  Length 622;
  Best Local Similarity   64.7%;  Pred. No. 0.0099;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTNP 17
           ||||    ||| |||:|
Db     158 RNPDSSTTGPWCYTTDP 174


RESULT 22
APOA_MACMU
ID   APOA_MACMU     STANDARD;      PRT;  1420 AA.
AC   P14417;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN   LPA.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89174660; PubMed=2925643;
RA   Tomlinson J.E., McLean J.W., Lawn R.M.;
RT   "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT   synthesis.";
RL   J. Biol. Chem. 264:5957-5965(1989).
CC   -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC       (Lp(a)). It has serine proteinase activity and is able of
CC       autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC       Lp(a) may be a ligand for megalin/Gp 330.
CC   -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC       decorin (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC       naturally occuring proteolytic fragments is correlated with
CC       atherosclerosis. Homology with plasminogen kringles IV and V is
CC       thought to underlie the atherogenicity of the protein, because the
CC       fragments are competing with plasminogen for fibrin(ogen) binding.
CC   -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC       leading to the formation of the so called mini-Lp(a). Apo(a)
CC       fragments accumulate in atherosclerotic lesions, where they may
CC       promote thrombogenesis. O-glycosylation may limit the extent of
CC       proteolytic fragmentation (By similarity).
CC   -!- SIMILARITY: CONTAINS AT LEAST 10 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04635; AAA36833.1; -.
DR   PIR; A30848; A30848.
DR   PIR; A32869; A32869.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.226; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 10.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00130; KR; 10.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 9.
DR   PROSITE; PS50070; KRINGLE_2; 10.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW   Kringle; Repeat; Atherosclerosis.
FT   NON_TER       1      1
FT   DOMAIN       49    127       KRINGLE 1.
FT   DOMAIN      163    241       KRINGLE 2.
FT   DOMAIN      277    355       KRINGLE 3.
FT   DOMAIN      391    469       KRINGLE 4.
FT   DOMAIN      505    583       KRINGLE 5.
FT   DOMAIN      619    697       KRINGLE 6.
```

```
FT   DOMAIN       725    803       KRINGLE 7.
FT   DOMAIN       839    917       KRINGLE 8.
FT   DOMAIN       953   1031       KRINGLE 9.
FT   DOMAIN      1067   1145       KRINGLE 10.
FT   DOMAIN      1191   1420       SERINE PROTEASE.
SQ   SEQUENCE   1420 AA;  158367 MW;  BE102949E03C5B0E CRC64;

  Query Match             47.5%;  Score 65.5;  DB 1;  Length 1420;
  Best Local Similarity   47.8%;  Pred. No. 0.15;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |||| ::  || || :||  ::|
Db     776 RNPDAEI-RPWCYTMDPRVRWEY 797


RESULT 23
YE90_MYCTU
ID   YE90_MYCTU     STANDARD;      PRT;   435 AA.
AC   P71771;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 48.2 kDa protein Rv1490.
GN   RV1490 OR MT1536 OR MTCY277.12.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z79701; CAB02040.1; -.
DR   EMBL; AE007022; AAK45803.1; -.
DR   TIGR; MT1536; -.
DR   TubercuList; Rv1490; -.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     40     60       POTENTIAL.
FT   TRANSMEM    103    123       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    195    215       POTENTIAL.
FT   TRANSMEM    226    246       POTENTIAL.
FT   TRANSMEM    313    333       POTENTIAL.
FT   TRANSMEM    358    378       POTENTIAL.
FT   TRANSMEM    381    401       POTENTIAL.
FT   TRANSMEM    414    434       POTENTIAL.
SQ   SEQUENCE   435 AA;  48149 MW;  1144B549CD418F30 CRC64;

  Query Match             43.5%;  Score 60;  DB 1;  Length 435;
  Best Local Similarity   47.6%;  Pred. No. 0.27;
  Matches   10;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       3 PDGDVGGPWAYTTNPRKLYDY 23
           | | |  |||  | |::: ||
Db     259 PHGMVENPWAVATTPQRVLDY 279


RESULT 24
ROR2_HUMAN
ID   ROR2_HUMAN     STANDARD;      PRT;   943 AA.
AC   Q01974; Q9HAY7; Q9HB61;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tyrosine-protein kinase transmembrane receptor ROR2 precursor
DE   (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
GN   ROR2 OR NTRKR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93100347; PubMed=1334494;
RA   Masiakowski P., Carroll R.D.;
RT   "A novel family of cell surface receptors with tyrosine kinase-like
RT   domain.";
RL   J. Biol. Chem. 267:26181-26190(1992).
RN   [2]
RP   SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
RX   MEDLINE=20164326; PubMed=10700182;
RA   Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,
RA   Pollitt C., DeChiara T.M., Kimble R.B., Valenzuela D.M.,
RA   Yancopoulos G.D., Wilkie A.O.M.;
RT   "Dominant mutations in ROR2, encoding an orphan receptor tyrosine
RT   kinase, cause brachydactyly type B.";
RL   Nat. Genet. 24:275-278(2000).
RN   [3]
RP   SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.
RX   MEDLINE=20442029; PubMed=10986040;
RA   Schwabe G.C., Tinschert S., Buschow C., Meinecke P., Wolff G.,
RA   Gillessen-Kaesbach G., Oldridge M., Wilkie A.O.M., Koemec R.,
RA   Mundlos S.;
RT   "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause
RT   brachydactyly type B.";
RL   Am. J. Hum. Genet. 67:822-831(2000).
RN   [4]
RP   VARIANTS RRS C-184; W-189; W-366 AND K-620.
RX   MEDLINE=20392394; PubMed=10932186;
RA   Afzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,
RA   Ternes-Pereira E., Tueysuez B., Murday V.A., Patton M.A.,
RA   Wilkie A.O.M., Jeffery S.;
RT   "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,
RT   is caused by mutation of ROR2.";
RL   Nat. Genet. 25:419-422(2000).
RN   [5]
RP   VARIANT RRS TYR-182.
RX   MEDLINE=20392395; PubMed=10932187;
RA   van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA   Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.;
RT   "Mutation of the gene encoding the ROR2 tyrosine kinase causes
RT   autosomal recessive Robinow syndrome.";
RL   Nat. Genet. 25:423-426(2000).
```

```
RN   [6]
RP   ERRATUM.
RA   van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA   Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.;
RL   Nat. Genet. 26:383-383(2000).
CC   -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED
CC       IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE
CC       REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
CC       EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
CC       DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES.
CC   -!- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF BRACHYDACTYLY TYPE B
CC       (BDB). BDB IS AN AUTOSOMAL DOMINANT SKELETAL DISORDER
CC       CHARACTERIZED BY HYPOPLASIA/APLASIA OF DISTAL PHALANGES AND NAILS.
CC       IN BDB THE MIDDLE PHALANGES ARE SHORT BUT IN ADDITION THE TERMINAL
CC       PHALANGES ARE RUDIMENTARY OR ABSENT. BOTH FINGERS AND TOES ARE
CC       AFFECTED. THE THUMBS AND BIG TOES ARE USUALLY DEFORMED.
CC   -!- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF RECESSIVE ROBINOW SYNDROME
CC       (RRS). RRS IS AN AUTOSOMAL DISORDER CHARACTERIZED BY SKELETAL
CC       DYSPLASIA WITH GENERALIZED LIMB BONE SHORTENING, SEGMENTAL DEFECTS
CC       OF THE SPINE, BRACHYDACTYLY AND A DYSMORPHIC FACIAL APPEARANCE.
CC   -!- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE ROR SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M97639; AAA60276.1; -.
DR   EMBL; AF294796; AAG01184.2; -.
DR   EMBL; AF254747; AAG01184.2; JOINED.
DR   EMBL; AF254748; AAG01184.2; JOINED.
DR   EMBL; AF254749; AAG01184.2; JOINED.
DR   EMBL; AF254750; AAG01184.2; JOINED.
DR   EMBL; AF254751; AAG01184.2; JOINED.
DR   EMBL; AF254752; AAG01184.2; JOINED.
DR   EMBL; AF254753; AAG01184.2; JOINED.
DR   EMBL; AF279762; AAG33132.1; -.
DR   EMBL; AF279755; AAG33132.1; JOINED.
DR   EMBL; AF279756; AAG33132.1; JOINED.
DR   EMBL; AF279757; AAG33132.1; JOINED.
DR   EMBL; AF279758; AAG33132.1; JOINED.
DR   EMBL; AF279759; AAG33132.1; JOINED.
DR   EMBL; AF279760; AAG33132.1; JOINED.
DR   EMBL; AF279761; AAG33132.1; JOINED.
DR   HSSP; P00747; 1KRN.
DR   MIM; 602337; -.
DR   MIM; 113000; -.
DR   MIM; 268310; -.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000024; Fz_domain.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF01392; Fz; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS50038; FZ; 1.
DR   PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW   Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW   Immunoglobulin domain; Developmental protein; Polymorphism;
KW   Disease mutation.
FT   SIGNAL        1     33       POTENTIAL.
FT   CHAIN        34    943       TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT                                 RECEPTOR ROR2.
FT   DOMAIN       34    403       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    404    424       POTENTIAL.
FT   DOMAIN      425    943       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       76    142       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      169    303       FZ.
FT   DOMAIN      316    394       KRINGLE.
FT   DOMAIN      473    746       PROTEIN KINASE.
FT   DOMAIN      753    782       SER/THR-RICH.
FT   DOMAIN      784    857       PRO-RICH.
FT   DOMAIN      859    882       SER/THR-RICH.
FT   NP_BIND     479    487       ATP (BY SIMILARITY).
FT   BINDING     507    507       ATP (BY SIMILARITY).
FT   ACT_SITE    615    615       BY SIMILARITY.
FT   MOD_RES     646    646       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   DISULFID     83    135       BY SIMILARITY.
FT   CARBOHYD     70     70       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     182    182       C -> Y (IN RRS).
FT                                 /FTId=VAR_010911.
FT   VARIANT     184    184       R -> C (IN RRS).
FT                                 /FTId=VAR_010768.
FT   VARIANT     189    189       R -> W (IN RRS).
FT                                 /FTId=VAR_010769.
FT   VARIANT     245    245       A -> T.
FT                                 /FTId=VAR_010912.
FT   VARIANT     366    366       R -> W (IN RRS).
FT                                 /FTId=VAR_010770.
FT   VARIANT     620    620       N -> K (IN RRS).
FT                                 /FTId=VAR_010771.
FT   VARIANT     819    819       V -> I.
FT                                 /FTId=VAR_010913.
SQ   SEQUENCE   943 AA;  104726 MW;  DBAC1E4622B5ECA0 CRC64;

  Query Match             42.8%;  Score 59;  DB 1;  Length 943;
  Best Local Similarity   56.2%;  Pred. No. 0.83;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTN 16
             ||| | : ||| :| |
Db     366 RNPGGQMEGPWCFTQN 381


RESULT 25
ROR2_MOUSE
ID   ROR2_MOUSE     STANDARD;      PRT;   944 AA.
AC   Q9Z138;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tyrosine-protein kinase transmembrane receptor ROR2 precursor
DE   (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2)
DE   (mROR2).
GN   ROR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=99248426; PubMed=10231392;
RA   Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
RA   Hatta T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.;
RT   "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT   mRor1, mRor2, during mouse development: implications in development
RT   and function of the nervous system.";
RL   Genes Cells 4:41-56(1999).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=20164325; PubMed=10700181;
RA   DeChiara T.M., Kimble R.B., Poueymirou W.T., Rojas J., Masiakowski P.,
RA   Valenzuela D.M., Yancopoulos G.D.;
RT   "Ror2, encoding a receptor-like tyrosine kinase, is required for
RT   cartilage and growth plate development.";
RL   Nat. Genet. 24:271-274(2000).
CC   -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED
CC       IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE
CC       REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE ROR SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB010384; BAA75481.1; -.
DR   HSSP; P00747; 5HPG.
DR   MGD; MGI:1347521; Ror2.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000024; Fz_domain.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF01392; Fz; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR00109; TYRKINASE.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS50038; FZ; 1.
DR   PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW   Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW   Immunoglobulin domain; Developmental protein.
FT   SIGNAL        1     33       POTENTIAL.
FT   CHAIN        34    944       TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT                                 RECEPTOR ROR2.
FT   DOMAIN       34    403       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    404    424       POTENTIAL.
FT   DOMAIN      425    944       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       76    142       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      169    303       FZ.
FT   DOMAIN      316    394       KRINGLE.
FT   DOMAIN      473    746       PROTEIN KINASE.
FT   DOMAIN      753    782       SER/THR-RICH.
FT   DOMAIN      784    857       PRO-RICH.
FT   DOMAIN      859    882       SER/THR-RICH.
FT   NP_BIND     479    487       ATP (BY SIMILARITY).
FT   BINDING     507    507       ATP (BY SIMILARITY).
FT   ACT_SITE    615    615       BY SIMILARITY.
FT   MOD_RES     646    646       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   DISULFID     83    135       BY SIMILARITY.
FT   CARBOHYD     70     70       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   944 AA;  105050 MW;  CD2EEBC710387A56 CRC64;

  Query Match             42.8%;  Score 59;  DB 1;  Length 944;
  Best Local Similarity   56.2%;  Pred. No. 0.83;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RNPDGDVGGPWAYTTN 16
           ||| | : ||| :| |
Db     366 RNPGGQMEGPWCFTQN 381


RESULT 26
TPA_MOUSE
ID   TPA_MOUSE      STANDARD;      PRT;   559 AA.
AC   P11214;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88087303; PubMed=2826484;
RA   Rickles R.J., Darrow A.L., Strickland S.;
RT   "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT   activator mRNA and its expression during F9 teratocarcinoma cell
RT   differentiation.";
RL   J. Biol. Chem. 263:1563-1569(1988).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03520; AAA40470.1; -.
```

```
DR   PIR; A29941; A29941.
DR   HSSP; P00750; 1TPG.
DR   MEROPS; S01.232; -.
DR   MGD; MGI:97610; Plat.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000083; fibronectin_type_1.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Plasma; Kringle; EGF-like domain; Signal.
FT   SIGNAL        1     17       PROBABLE.
FT   PROPEP       18     29
FT   CHAIN        30    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN        30    308       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                                CHAIN.
FT   CHAIN       309    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                                CHAIN.
FT   DOMAIN       36     78       FIBRONECTIN TYPE-I.
FT   DOMAIN       79    117       EGF-LIKE.
FT   DOMAIN      124    205       KRINGLE 1.
FT   DOMAIN      213    294       KRINGLE 2.
FT   DOMAIN      309    559       SERINE PROTEASE.
FT   ACT_SITE    355    355       CHARGE RELAY SYSTEM.
FT   ACT_SITE    404    404       CHARGE RELAY SYSTEM.
FT   ACT_SITE    510    510       CHARGE RELAY SYSTEM.
FT   DISULFID     38     68       BY SIMILARITY.
FT   DISULFID     66     75       BY SIMILARITY.
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID     88    105       BY SIMILARITY.
FT   DISULFID    107    116       BY SIMILARITY.
FT   DISULFID    124    205       BY SIMILARITY.
FT   DISULFID    145    187       BY SIMILARITY.
FT   DISULFID    176    200       BY SIMILARITY.
FT   DISULFID    213    294       BY SIMILARITY.
FT   DISULFID    234    276       BY SIMILARITY.
FT   DISULFID    265    289       BY SIMILARITY.
FT   DISULFID    297    428       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    340    356       BY SIMILARITY.
FT   DISULFID    348    417       BY SIMILARITY.
FT   DISULFID    442    516       BY SIMILARITY.
FT   DISULFID    474    490       BY SIMILARITY.
FT   DISULFID    506    534       BY SIMILARITY.
FT   CARBOHYD    149    149       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    481    481       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   559 AA;  63110 MW;  4ACEE57DC6A282A5 CRC64;

  Query Match             40.6%;  Score 56;  DB 1;  Length 559;
  Best Local Similarity   50.0%;  Pred. No. 1.3;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 RNPDGDVGGPWAYTTNPRKL-YDY 23
           ||||||   || :    ||| ::|
Db     266 RNPDGD-ARPWCHVMKDRKLTWEY 288
```

```
RESULT 27
TPA_RAT
ID   TPA_RAT        STANDARD;      PRT;   559 AA.
AC   P19637;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89170114; PubMed=3148445;
RA   Ny T., Leonardsson G., Hsueh A.J.W.;
RT   "Cloning and characterization of a cDNA for rat tissue-type
RT   plasminogen activator.";
RL   DNA 7:671-677(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90130448; PubMed=2105315;
RA   Feng P., Ohlsson M., Ny T.;
RT   "The structure of the TATA-less rat tissue-type plasminogen activator
RT   gene. Species-specific sequence divergences in the promoter predict
RT   differences in regulation of gene expression.";
RL   J. Biol. Chem. 265:2022-2027(1990).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M23697; AAA41812.1; -.
DR   EMBL; M31197; AAA42261.1; -.
DR   EMBL; M31185; AAA42261.1; JOINED.
DR   EMBL; M31186; AAA42261.1; JOINED.
DR   EMBL; M31187; AAA42261.1; JOINED.
DR   EMBL; M31188; AAA42261.1; JOINED.
DR   EMBL; M31189; AAA42261.1; JOINED.
DR   EMBL; M31190; AAA42261.1; JOINED.
DR   EMBL; M31191; AAA42261.1; JOINED.
DR   EMBL; M31192; AAA42261.1; JOINED.
DR   EMBL; M31193; AAA42261.1; JOINED.
DR   EMBL; M31194; AAA42261.1; JOINED.
DR   EMBL; M31195; AAA42261.1; JOINED.
DR   EMBL; M31196; AAA42261.1; JOINED.
```

```
DR   EMBL; A19618; CAA01482.1; -.
DR   PIR; A31597; A31597.
DR   PIR; A35029; A35029.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000083; fibronectin_type_1.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Plasma; Kringle; EGF-like domain; Signal.
FT   SIGNAL       1     17       PROBABLE.
FT   PROPEP      18     29
FT   CHAIN       30    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN       30    308       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                               CHAIN.
FT   CHAIN      309    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                               CHAIN.
FT   DOMAIN      36     78       FIBRONECTIN TYPE-I.
FT   DOMAIN      79    117       EGF-LIKE.
FT   DOMAIN     124    205       KRINGLE 1.
FT   DOMAIN     213    294       KRINGLE 2.
FT   DOMAIN     309    559       SERINE PROTEASE.
FT   ACT_SITE   355    355       CHARGE RELAY SYSTEM.
FT   ACT_SITE   404    404       CHARGE RELAY SYSTEM.
FT   ACT_SITE   510    510       CHARGE RELAY SYSTEM.
FT   DISULFID    38     68       BY SIMILARITY.
FT   DISULFID    66     75       BY SIMILARITY.
FT   DISULFID    83     94       BY SIMILARITY.
FT   DISULFID    88    105       BY SIMILARITY.
FT   DISULFID   107    116       BY SIMILARITY.
FT   DISULFID   124    205       BY SIMILARITY.
FT   DISULFID   145    187       BY SIMILARITY.
FT   DISULFID   176    200       BY SIMILARITY.
FT   DISULFID   213    294       BY SIMILARITY.
FT   DISULFID   234    276       BY SIMILARITY.
FT   DISULFID   265    289       BY SIMILARITY.
FT   DISULFID   297    428       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   340    356       BY SIMILARITY.
FT   DISULFID   348    417       BY SIMILARITY.
FT   DISULFID   442    516       BY SIMILARITY.
FT   DISULFID   474    490       BY SIMILARITY.
FT   DISULFID   506    534       BY SIMILARITY.
FT   CARBOHYD   149    149       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   481    481       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   380    380       E -> K (IN REF. 1).
SQ   SEQUENCE   559 AA;  62903 MW;  7DBD3809C1D1C921 CRC64;

  Query Match             40.6%;  Score 56;  DB 1;  Length 559;
  Best Local Similarity   50.0%;  Pred. No. 1.3;
  Matches   12;  Conservative    3;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 RNPDGDVGGPWAYTTNPRKL-YDY 23
           ||||||   || :    ||| ::|
Db     266 RNPDGD-AKPWCHVMKDRKLTWEY 288


RESULT 28
PHLN_BURPS
ID   PHLN_BURPS     STANDARD;      PRT;   700 AA.
AC   Q9RGS8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Non-hemolytic phospholipase C precursor (EC 3.1.4.3) (PLC-N)
DE   (Phosphatidylcholine cholinephosphohydrolase) (Phosphatidylcholine-
DE   hydrolyzing phospholipase C) (PC-PLC).
GN   PLCN.
OS   Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=28450;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=SPT1;
RX   MEDLINE=99454884; PubMed=10523590;
RA   Korbsrisate S., Suwanasai N., Leelaporn A., Ezaki T., Kawamura Y.,
RA   Sarasombath S.;
RT   "Cloning and characterization of a nonhemolytic phospholipase C gene
RT   from Burkholderia pseudomallei.";
RL   J. Clin. Microbiol. 37:3742-3745(1999).
CC   -!- FUNCTION: HYDROLYZES PHOSPHATIDYLSERINE AS WELL AS
CC       PHOSPHATIDYLCHOLINE.
CC   -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC       diacylglycerol + choline phosphate.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL PHOSPHOLIPASE C FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF107252; AAF17299.1; -.
KW   Hydrolase; Signal.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    700       NON-HEMOLYTIC PHOSPHOLIPASE C.
SQ   SEQUENCE   700 AA;  77190 MW;  5EDAE6FCD0FB129B CRC64;

  Query Match             38.0%;  Score 52.5;  DB 1;  Length 700;
  Best Local Similarity   41.4%;  Pred. No. 5.3;
  Matches   12;  Conservative    0;  Mismatches    4;  Indels   13;  Gaps    1;

Qy       7 VGGPWAYTTNPRK----------LYD 22
           | ||||||   ||           |||
Db     548 VDGPWAYTVEARKRLSDEWSTALTLSLYD 576


RESULT 29
TPA_HUMAN
ID   TPA_HUMAN      STANDARD;      PRT;   562 AA.
AC   P00750; Q15103;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).
GN   PLAT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Melanoma;
```

RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friezner Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [9]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1368681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [10]
RP SEQUENCE OF 36-562.
RC TISSUE=Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [11]
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE=Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator.";
RL Eur. J. Biochem. 132:681-686(1983).
RN [12]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [13]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [14]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT threonine-61 in the epidermal growth factor domain.";
RL Biochemistry 30:2311-2314(1991).
RN [15]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244765; PubMed=1645336;
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in Escherichia coli.";
RL J. Biol. Chem. 266:10070-10072(1991).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 A crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator.";
RL J. Mol. Biol. 258:117-135(1996).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; PubMed=1310033;
RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinskly A.,
RA Westbrook M.L., Kossiakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
RT activator at 2.4-A resolution.";
RL Biochemistry 31:270-279(1992).
RN [19]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; PubMed=2558718;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
RT from human tissue-type plasminogen activator.";
RL Biochemistry 28:9350-9360(1989).
RN [20]

```
RP   STRUCTURE BY NMR OF KRINGLE 2.
RX   MEDLINE=91200042; PubMed=1901789;
RA   Byeon I.-J.L., Kelley R.F., Llinas M.;
RT   "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT   assignments and secondary structure.";
RL   Eur. J. Biochem. 197:155-165(1991).
RN   [21]
RP   STRUCTURE BY NMR OF KRINGLE 2.
RX   MEDLINE=92106329; PubMed=1762144;
RA   Byeon I.-J.L., Llinas M.;
RT   "Solution structure of the tissue-type plasminogen activator kringle
RT   2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT   drug.";
RL   J. Mol. Biol. 222:1035-1051(1991).
RN   [22]
RP   STRUCTURE BY NMR OF 38-85.
RX   MEDLINE=92292163; PubMed=1602484;
RA   Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
RA   Baron M., Campbell I.D.;
RT   "Solution structure of the fibrin binding finger domain of
RT   tissue-type plasminogen activator determined by 1H nuclear magnetic
RT   resonance.";
RL   J. Mol. Biol. 225:821-833(1992).
RN   [23]
RP   STRUCTURE BY NMR OF 36-126.
RX   MEDLINE=96027104; PubMed=7582899;
RA   Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
RT   "The solution structure and backbone dynamics of the fibronectin type
RT   I and epidermal growth factor-like pair of modules of tissue-type
RT   plasminogen activator.";
RL   Structure 3:823-833(1995).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

  Query Match             37.7%;  Score 52;  DB 1;  Length 562;
  Best Local Similarity   45.8%;  Pred. No. 5;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 RNPDGDVGGPWAYTTNPRKL-YDY 23
           ||||||   || :    |:| ::|
Db     268 RNPDGD-AKPWCHVLKNRRLTWEY 290


RESULT 30
Y4HQ_RHISN
ID   Y4HQ_RHISN     STANDARD;      PRT;   194 AA.
AC   P50361;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 21.7 kDa protein Y4HQ.
GN   Y4HQ.
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
RN   [2]
RP   SEQUENCE OF 1-52 FROM N.A.
RA   Rochepeau P., Fellay R., Broughton W.J.;
RL   Submitted (JUL-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO A.RHIZOGENES PLASMID PRIA4B ORF-3 IN VIRA REGION.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74068; CAA52198.1; ALT_INIT.
DR   EMBL; AE000077; AAB92455.1; -.
KW   Hypothetical protein; Plasmid.
FT   CONFLICT      18     18       I -> M (IN REF. 2).
FT   CONFLICT      30     30       R -> S (IN REF. 2).
SQ   SEQUENCE   194 AA;  21747 MW;  647B7161AF1513E3 CRC64;

  Query Match             37.0%;  Score 51;  DB 1;  Length 194;
  Best Local Similarity   72.7%;  Pred. No. 2.3;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 PDGDVGGPWAY 13
           |  |||||| |
Db     128 PPEDVGGPWGY 138


RESULT 31
EST1_SCHGA
ID   EST1_SCHGA     STANDARD;      PRT;   198 AA.
AC   P81429;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Esterase SG1 precursor (EC 3.1.1.1) (Carboxylic-ester hydrolase)
DE   (Fragment).
GN   SG1.
OS   Schizaphis graminum (Aphid).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
OC   Aphidiformes; Aphidoidea; Aphididae; Aphidini; Schizaphis.
OX   NCBI_TaxID=13262;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 22-41.
RX   MEDLINE=20078165; PubMed=10612041;
RA   Ono M., Swanson J.J., Field L.M., Devonshire A.L., Siegfried B.D.;
RT   "Amplification and methylation of an esterase gene associated with
RT   insecticide-resistance in greenbugs, Schizaphis graminum.";
RL   Insect Biochem. Mol. Biol. 29:1065-1073(1999).
CC   -!- FUNCTION: OVERPRODUCTION OF NONSPECIFIC ESTERASES IS A COMMON
CC       MECHANISM OF RESISTANCE TO ORGANOPHOSPHATE INSECTICIDES.
CC   -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC       carboxylic anion.
CC   -!- MISCELLANEOUS: THIS ESTERASE CONFERS OP INSECTICIDE RESISTANCE.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y17580; CAA76792.1; -.
DR   HSSP; P37967; 1QE3.
DR   InterPro; IPR002018; Carboxylesterase_B.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; PARTIAL.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL        1     21
FT   CHAIN        22   >198       ESTERASE SG1.
```

```
FT   DISULFID     87    104       BY SIMILARITY.
FT   CARBOHYD     79     79       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     24     24       A -> T (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     28     28       R -> N (IN REF. 1; AA SEQUENCE).
FT   NON_TER     198    198
SQ   SEQUENCE   198 AA;  21897 MW;  EFEF43EA6A5D5987 CRC64;

  Query Match             37.0%;  Score 51;  DB 1;  Length 198;
  Best Local Similarity   42.1%;  Pred. No. 2.3;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 GDVGGPWAYTTNPRKLYDY 23
           ||: | : || : ||:| :
Db      32 GDIAGGYEYTYSGRKIYSF 50


RESULT 32
ESTE_MYZPE
ID   ESTE_MYZPE     STANDARD;      PRT;   552 AA.
AC   P35501;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   Esterase E4 precursor (EC 3.1.1.1) (Carboxylic-ester hydrolase).
OS   Myzus persicae (Peach-potato aphid).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
OC   Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Myzus.
OX   NCBI_TaxID=13164;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-63.
RC   STRAIN=R3 / ISOLATE 794J;
RX   MEDLINE=93384534; PubMed=8373371;
RA   Field L.M., Williamson M.S., Moores G.D., Devonshire A.L.;
RT   "Cloning and analysis of the esterase genes conferring insecticide
RT   resistance in the peach-potato aphid, Myzus persicae (Sulzer).";
RL   Biochem. J. 294:569-574(1993).
CC   -!- FUNCTION: OVERPRODUCTION OF NONSPECIFIC ESTERASES IS A COMMON
CC       MECHANISM OF RESISTANCE TO ORGANOPHOSPHATE INSECTICIDES.
CC   -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC       carboxylic anion.
CC   -!- MISCELLANEOUS: THIS ESTERASE CONFERS INSECTICIDE RESISTANCE.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74554; CAA52648.1; -.
DR   PIR; S36786; S36786.
DR   HSSP; P21836; 1MAH.
DR   InterPro; IPR002018; Carboxylesterase_B.
DR   InterPro; IPR000379; Est_lip_thioest_actsite.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; FALSE_NEG.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    552       ESTERASE E4.
FT   ACT_SITE    214    214       BY SIMILARITY.
FT   ACT_SITE    339    339       BY SIMILARITY.
FT   ACT_SITE    463    463       BY SIMILARITY.
FT   DISULFID     89    106       BY SIMILARITY.
FT   DISULFID    266    277       BY SIMILARITY.
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    371    371       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    404    404       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    443    443       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   552 AA;  61348 MW;  B97B67272DFF7209 CRC64;

  Query Match             37.0%;  Score 51;  DB 1;  Length 552;
  Best Local Similarity   42.1%;  Pred. No. 6.8;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 GDVGGPWAYTTNPRKLYDY 23
           |:: | : || | ||:| :
Db      34 GEIAGGFEYTYNGRKIYSF 52


RESULT 33
ESTF_MYZPE
ID   ESTF_MYZPE     STANDARD;      PRT;   564 AA.
AC   P35502;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   Esterase FE4 precursor (EC 3.1.1.1) (Carboxylic-ester hydrolase).
OS   Myzus persicae (Peach-potato aphid).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
OC   Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Myzus.
OX   NCBI_TaxID=13164;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-63.
RC   STRAIN=ISOLATE 800F;
RX   MEDLINE=93384534; PubMed=8373371;
RA   Field L.M., Williamson M.S., Moores G.D., Devonshire A.L.;
RT   "Cloning and analysis of the esterase genes conferring insecticide
RT   resistance in the peach-potato aphid, Myzus persicae (Sulzer).";
RL   Biochem. J. 294:569-574(1993).
CC   -!- FUNCTION: OVERPRODUCTION OF NONSPECIFIC ESTERASES IS A COMMON
CC       MECHANISM OF RESISTANCE TO ORGANOPHOSPHATE INSECTICIDES.
CC   -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC       carboxylic anion.
CC   -!- MISCELLANEOUS: THIS ESTERASE CONFERS INSECTICIDE RESISTANCE.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74555; CAA52649.1; -.
DR   PIR; S36787; S36787.
DR   HSSP; P21836; 1MAA.
DR   InterPro; IPR002018; Carboxylesterase_B.
DR   InterPro; IPR000379; Est_lip_thioest_actsite.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; FALSE_NEG.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    564       ESTERASE FE4.
FT   ACT_SITE    214    214       BY SIMILARITY.
FT   ACT_SITE    339    339       BY SIMILARITY.
FT   ACT_SITE    463    463       BY SIMILARITY.
FT   DISULFID     89    106       BY SIMILARITY.
FT   DISULFID    266    277       BY SIMILARITY.
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    371    371       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    404    404       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    443    443       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   564 AA;  62757 MW;  128CC0EC3F2D9F1A CRC64;

  Query Match             37.0%;  Score 51;  DB 1;  Length 564;
  Best Local Similarity   42.1%;  Pred. No. 6.9;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy       5 GDVGGPWAYTTNPRKLYDY 23
             |:: | : || | ||:| :
Db      34 GEIAGGFEYTYNGRKIYSF 52


RESULT 34
TPA_BOVIN
ID   TPA_BOVIN      STANDARD;      PRT;   566 AA.
AC   Q28198;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Ravn P., Berglund L., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen activators uPA
RT   and tPA.";
RL   Int. Dairy J. 5:605-617(1995).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X85800; CAA59795.1; -.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000083; fibronectin_type_1.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Plasma; Kringle; EGF-like domain; Signal.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   PROPEP       22     33       BY SIMILARITY.
FT   CHAIN        34    566       TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN        34    314       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                                CHAIN.
FT   CHAIN       315    566       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                                CHAIN.
FT   DOMAIN       40     82       FIBRONECTIN TYPE-I.
FT   DOMAIN       83    121       EGF-LIKE.
FT   DOMAIN      128    209       KRINGLE 1.
FT   DOMAIN      219    300       KRINGLE 2.
FT   DOMAIN      315    566       SERINE PROTEASE.
FT   ACT_SITE    361    361       CHARGE RELAY SYSTEM.
FT   ACT_SITE    410    410       CHARGE RELAY SYSTEM.
FT   ACT_SITE    517    517       CHARGE RELAY SYSTEM.
FT   DISULFID     42     72       BY SIMILARITY.
FT   DISULFID     70     79       BY SIMILARITY.
FT   DISULFID     87     98       BY SIMILARITY.
FT   DISULFID     92    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    128    209       BY SIMILARITY.
FT   DISULFID    149    191       BY SIMILARITY.
FT   DISULFID    180    204       BY SIMILARITY.
FT   DISULFID    219    300       BY SIMILARITY.
FT   DISULFID    240    282       BY SIMILARITY.
FT   DISULFID    271    295       BY SIMILARITY.
FT   DISULFID    303    434       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    346    362       BY SIMILARITY.
FT   DISULFID    354    423       BY SIMILARITY.
FT   DISULFID    448    523       BY SIMILARITY.
FT   DISULFID    480    496       BY SIMILARITY.
FT   DISULFID    513    541       BY SIMILARITY.
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    487    487       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   566 AA;  63701 MW;  2EB6BEB4E32276C3 CRC64;

  Query Match             37.0%;  Score 51;  DB 1;  Length 566;
  Best Local Similarity   45.8%;  Pred. No. 7;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 RNPDGDVGGPWAYTTNPRKL-YDY 23
             ||||||   || :    |:| ::|
Db     272 RNPDGD-AQPWCHVWKDRQLTWEY 294


RESULT 35
URTG_DESRO
ID   URTG_DESRO     STANDARD;      PRT;   394 AA.
AC   P49150;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE   gamma).
OS   Desmodus rotundus (Vampire bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC   Desmodontinae; Desmodus.
OX   NCBI_TaxID=9430;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Salivary gland;
RX   MEDLINE=92039036; PubMed=1937019;
RA   Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA   Alagon A., Donner P., Schleuning W.D.;
RT   "The plasminogen activator family from the salivary gland of the
RT   vampire bat Desmodus rotundus: cloning and expression.";
RL   Gene 105:229-237(1991).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=93393059; PubMed=1309059;
RA   Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA   Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA   Donner P.;
RT   "Plasminogen activators from the saliva of Desmodus rotundus (common
RT   vampire bat): unique fibrin specificity.";
RL   Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC   -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC       EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC       AGENT.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M63990; AAA31595.1; -.
DR   HSSP; P98119; 1A5I.
DR   MEROPS; S01.239; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; Signal; Multigene family.
FT   SIGNAL        1     36       POTENTIAL.
FT   CHAIN        37    394       SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT   DOMAIN       45    126       KRINGLE.
FT   DOMAIN      142    394       SERINE PROTEASE.
FT   ACT_SITE    189    189       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    238    238       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    345    345       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     45    126       BY SIMILARITY.
FT   DISULFID     66    108       BY SIMILARITY.
FT   DISULFID     97    121       BY SIMILARITY.
FT   DISULFID    131    262       BY SIMILARITY.
FT   DISULFID    174    190       BY SIMILARITY.
FT   DISULFID    182    251       BY SIMILARITY.
FT   DISULFID    276    351       BY SIMILARITY.
FT   DISULFID    308    324       BY SIMILARITY.
FT   DISULFID    341    369       BY SIMILARITY.
FT   CARBOHYD    315    315       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   394 AA;  44105 MW;  9CCD6F52F3D81FCD CRC64;

  Query Match             36.6%;  Score 50.5;  DB 1;  Length 394;
  Best Local Similarity   52.6%;  Pred. No. 5.6;
  Matches   10;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRK 19
           |||||    || |    ||
Db      98 RNPDG-ASKPWCYVIKARK 115


RESULT 36
FA12_HUMAN
ID   FA12_HUMAN     STANDARD;      PRT;   615 AA.
AC   P00748; P78339;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE   (HAF).
GN   F12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88007593; PubMed=2888762;
RA   Cool D.E., McGillivray R.T.A.;
RT   "Characterization of the human blood coagulation factor XII gene.
RT   Intron/exon gene organization and analysis of the 5'-flanking
RT   region.";
RL   J. Biol. Chem. 262:13662-13673(1987).
RN   [2]
RP   SEQUENCE OF 4-615 FROM N.A.
RX   MEDLINE=86176794; PubMed=3754331;
RA   Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA   Cortese R.;
RT   "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL   Nucleic Acids Res. 14:3146-3146(1986).
RN   [3]
RP   SEQUENCE OF 14-615 FROM N.A.
RX   MEDLINE=86033830; PubMed=3877053;
RA   Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA   McGillivray R.T.A.;
RT   "Characterization of human blood coagulation factor XII cDNA.
RT   Prediction of the primary structure of factor XII and the tertiary
RT   structure of beta-factor XIIa.";
RL   J. Biol. Chem. 260:13666-13676(1985).
RN   [4]
RP   SEQUENCE OF 146-615 FROM N.A.
RX   MEDLINE=86216049; PubMed=3011063;
RA   Que B.G., Davie E.W.;
RT   "Characterization of a cDNA coding for human factor XII (Hageman
RT   factor).";
RL   Biochemistry 25:1525-1528(1986).
RN   [5]
RP   SEQUENCE OF 20-379.
RX   MEDLINE=85182674; PubMed=3886654;
RA   McMullen B.A., Fujikawa K.;
RT   "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT   (activated Hageman factor).";
RL   J. Biol. Chem. 260:5328-5341(1985).
RN   [6]
RP   SEQUENCE OF 354-362 AND 373-615.
RX   MEDLINE=83291041; PubMed=6604055;
RA   Fujikawa K., McMullen B.A.;
RT   "Amino acid sequence of human beta-factor XIIa.";
RL   J. Biol. Chem. 258:10924-10933(1983).
RN   [7]
RP   SEQUENCE OF 561-615 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=96133302; PubMed=8528215;
RA   Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT   "The novel acceptor splice site mutation 11396(G-->A) in the factor
RT   XII gene causes a truncated transcript in cross-reacting material
```

```
RT   negative patients.";
RL   Hum. Mol. Genet. 4:1235-1237(1995).
RN   [8]
RP   VARIANT WASHINGTON DC.
RX   MEDLINE=90046788; PubMed=2510163;
RA   Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA   Saito H.;
RT   "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT   factor XIIa results from Cys-571-->Ser substitution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN   [9]
RP   VARIANT LOCARNO.
RX   MEDLINE=94325559; PubMed=8049433;
RA   Hovinga J.K., Schaller J., Stricker H., Wuillemin W.A., Furlan M.,
RA   Laemmle B.;
RT   "Coagulation factor XII Locarno: the functional defect is caused by
RT   the amino acid substitution Arg-353-->Pro leading to loss of a
RT   kallikrein cleavage site.";
RL   Blood 84:1173-1181(1994).
RN   [10]
RP   CARBOHYDRATE-LINKAGE SITE THR-109.
RX   MEDLINE=92184750; PubMed=1544894;
RA   Harris R.J., Ling V.T., Spellman M.W.;
RT   "O-linked fucose is present in the first epidermal growth factor
RT   domain of factor XII but not protein C.";
RL   J. Biol. Chem. 267:5102-5107(1992).
CC   -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC       THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC       GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC   -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC       VII to form factor VIIa and factor XI to form factor XIa.
CC   -!- PTM: O- AND N-GLYCOSYLATED.
CC   -!- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC       SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC   -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC       COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC       FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC       TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC       XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M31315; AAA70225.1; -.
DR   EMBL; M11723; AAA51986.1; -.
DR   EMBL; M17466; AAB59490.1; -.
DR   EMBL; M17464; AAB59490.1; JOINED.
DR   EMBL; M17465; AAB59490.1; JOINED.
DR   EMBL; M13147; AAA70224.1; -.
DR   EMBL; U71274; AAB51203.1; -.
DR   PIR; A29411; KFHU12.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.211; -.
DR   MIM; 234000; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000083; fibronectin_type_1.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00013; FNTYPEII.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW   Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW   Disease mutation.
FT   SIGNAL       1     19
FT   CHAIN       20    372       ALPHA-FACTOR XIIA HEAVY CHAIN.
FT   CHAIN      373    615       ALPHA-FACTOR XIIA LIGHT CHAIN.
FT   CHAIN      354    362       BETA-FACTOR XIIA PART 1.
FT   CHAIN      373    615       BETA-FACTOR XIIA PART 2.
FT   DOMAIN      47     88       FIBRONECTIN TYPE-II.
FT   DOMAIN      94    131       EGF-LIKE 1.
FT   DOMAIN     133    173       FIBRONECTIN TYPE-I.
FT   DOMAIN     174    210       EGF-LIKE 2.
FT   DOMAIN     217    295       KRINGLE.
FT   DOMAIN     296    349       PRO-RICH.
FT   DOMAIN     373    615       SERINE PROTEASE.
FT   CARBOHYD   109    109       O-LINKED (FUC).
FT   CARBOHYD   249    249       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   299    299       O-LINKED (POTENTIAL).
FT   CARBOHYD   305    305       O-LINKED (POTENTIAL).
FT   CARBOHYD   308    308       O-LINKED (POTENTIAL).
FT   CARBOHYD   328    328       O-LINKED (POTENTIAL).
FT   CARBOHYD   329    329       O-LINKED (POTENTIAL).
FT   CARBOHYD   337    337       O-LINKED (POTENTIAL).
FT   ACT_SITE   412    412       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   461    461       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   563    563       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    98    110       BY SIMILARITY.
FT   DISULFID   104    119       BY SIMILARITY.
FT   DISULFID   121    130       BY SIMILARITY.
FT   DISULFID   135    163       BY SIMILARITY.
FT   DISULFID   161    170       BY SIMILARITY.
FT   DISULFID   178    189       BY SIMILARITY.
FT   DISULFID   183    198       BY SIMILARITY.
FT   DISULFID   200    209       BY SIMILARITY.
FT   DISULFID   217    295       BY SIMILARITY.
FT   DISULFID   238    277       BY SIMILARITY.
FT   DISULFID   266    290       BY SIMILARITY.
FT   DISULFID   359    486       BY SIMILARITY.
FT   DISULFID   397    413       BY SIMILARITY.
FT   DISULFID   405    475       BY SIMILARITY.
FT   DISULFID   436    439       BY SIMILARITY.
FT   DISULFID   500    569       BY SIMILARITY.
FT   DISULFID   532    548       BY SIMILARITY.

  Query Match             35.5%;  Score 49;  DB 1;  Length 615;
  Best Local Similarity   41.7%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 RNPDGDVGGPWAYTTNPRKL-YDY 23
           |||| |:  || :  |  :| ::|
Db     267 RNPDNDI-RPWCFVLNRDRLSWEY 289
```

```
RESULT 37
PMA7_ARATH
ID   PMA7_ARATH     STANDARD;      PRT;   961 AA.
AC   Q9LY32;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   ATPase 7, plasma membrane-type (EC 3.6.3.6) (Proton pump 7).
GN   AHA7 OR AT3G60330 OR F27H5_120.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
CC   -!- FUNCTION: THE PLASMA MEMBRANE H+ ATPASE OF PLANTS AND FUNGI
CC       GENERATES A PROTON GRADIENT THAT DRIVES THE ACTIVE TRANSPORT OF
CC       NUTRIENTS BY H+-SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION
CC       AND/OR INTERNAL ALKINIZATION MAY MEDIATE GROWTH RESPONSES (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(IN) = ADP + PHOSPHATE +
CC       H(+)(OUT).
CC   -!- SUBUNIT: BINDS TO 14-3-3 PROTEINS. THE BINDING IS INDUCED BY
CC       PHOSPHORYLATION OF THR-960. BINDING TO 14-3-3 PROTEINS ACTIVATES
CC       THE H+-ATPASE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC       (E1-E2 ATPASES). SUBFAMILY IIIA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL163852; CAB87870.1; -.
DR   InterPro; IPR004014; Cation_ATPase.
DR   InterPro; IPR001757; E1-E2_ATPase.
DR   InterPro; IPR000695; HATPase.
DR   InterPro; IPR001454; Hydrolase.
DR   Pfam; PF00690; Cation_ATPase_N; 1.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PRINTS; PR00120; HATPASE.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Hydrolase; Hydrogen ion transport; Transmembrane; Phosphorylation;
KW   ATP-binding; Metal-binding; Magnesium; Multigene family;
KW   Hypothetical protein.
FT   DOMAIN        1     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     84       1 (POTENTIAL).
FT   DOMAIN       85     96       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     97    117       2 (POTENTIAL).
FT   DOMAIN      118    246       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    247    267       3 (POTENTIAL).
FT   DOMAIN      268    276       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    277    294       4 (POTENTIAL).
FT   DOMAIN      295    643       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    644    665       5 (POTENTIAL).
FT   DOMAIN      666    670       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    671    693       6 (POTENTIAL).
FT   DOMAIN      694    709       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    710    730       7 (POTENTIAL).
FT   DOMAIN      731    764       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    765    785       8 (POTENTIAL).
FT   DOMAIN      786    797       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    798    818       9 (POTENTIAL).
FT   DOMAIN      819    826       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    827    847       10 (POTENTIAL).
FT   DOMAIN      848    961       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES     332    332       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     960    960       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL       588    588       MAGNESIUM (BY SIMILARITY).
FT   METAL       592    592       MAGNESIUM (BY SIMILARITY).
FT   SITE        959    961       INTERACTION WITH 14-3-3 PROTEINS (BY
FT                                 SIMILARITY).
SQ   SEQUENCE   961 AA;  105520 MW;  BF8F9C5BC234A4F8 CRC64;

  Query Match             35.1%;  Score 48.5;  DB 1;  Length 961;
  Best Local Similarity   64.3%;  Pred. No. 28;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       3 PDGDV---GGPWAY 13
           |||||   |||| :
Db     469 PDGDVKGEGGPWDF 482


RESULT 38
Y145_METJA
ID   Y145_METJA     STANDARD;      PRT;   268 AA.
AC   Q57609;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MJ0145.
GN   MJ0145.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
```

```
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U67471; AAB98128.1; -.
DR   TIGR; MJ0145; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   268 AA;  30285 MW;  627BBD523868B05B CRC64;

  Query Match             34.8%;  Score 48;  DB 1;  Length 268;
  Best Local Similarity   57.1%;  Pred. No. 8.7;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       9 GPWAYTTNPRKLYD 22
           |||  | |||:  |
Db      13 GPWTVTPNPRRESD 26


RESULT 39
PHLC_MYCTU
ID   PHLC_MYCTU     STANDARD;      PRT;   517 AA.
AC   P95245;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phospholipase C 3 precursor (EC 3.1.4.3).
GN   PLCC OR RV2349C OR MT2414 OR MTCY98.18C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC       diacylglycerol + choline phosphate.
CC   -!- MISCELLANEOUS: POLYMORPHISM WAS DISCOVERED IN THE PHOSPHOLIPASE
CC       PLCA/B/C REGION.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL PHOSPHOLIPASE C FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z83860; CAB06146.1; ALT_INIT.
DR   EMBL; AE007081; AAK46707.1; ALT_INIT.
DR   TIGR; MT2414; -.
DR   TubercuList; Rv2349c; -.
KW   Hydrolase; Signal; Complete proteome.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    517       PHOSPHOLIPASE C 3.
SQ   SEQUENCE   517 AA;  55896 MW;  5B9D4241E09755AE CRC64;

  Query Match             34.8%;  Score 48;  DB 1;  Length 517;
  Best Local Similarity   88.9%;  Pred. No. 17;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 NPDGDVGGGP 10
           ||||| |||
Db     205 NPDGDQGGGP 213


RESULT 40
GAG_SIVA1
ID   GAG_SIVA1      STANDARD;      PRT;   520 AA.
AC   P27972;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   GAG polyprotein [Contains: Core proteins P17, P24, and P15].
GN   GAG.
OS   Simian immunodeficiency virus (AGM155 isolate) (SIV-AGM).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11727;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90156504; PubMed=2304139;
RA   Johnson P.R., Fomsgaard A., Allan J., Gravell M., London W.T.,
RA   Olmstead R.A., Hirsch V.M.;
RT   "Simian immunodeficiency viruses from African green monkeys display
RT   unusual genetic diversity.";
RL   J. Virol. 64:1086-1092(1990).
CC   -!- MISCELLANEOUS: THE 155 ISOLATE IS FROM A MONKEY IMPORTED FROM
CC       KENYA.
CC   -!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M29975; AAA91905.1; -.
DR   HSSP; P03351; 1EIA.
DR   InterPro; IPR000721; Gag_p24.
DR   InterPro; IPR000071; Retroviral_gag_p17.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00540; gag_p17; 1.
DR   Pfam; PF00607; gag_p24; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00234; HIV1MATRIX.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   AIDS; Core protein; Polyprotein; Zinc-finger; Repeat.
FT   CHAIN         1    141       CORE PROTEIN P17.
FT   CHAIN       142    520       CORE PROTEINS P24 AND P15.
FT   ZN_FING     398    415       CCHC-TYPE 1.
```

```
FT   ZN_FING     419    436       CCHC-TYPE 2.
SQ   SEQUENCE   520 AA;  57735 MW;  2FE1B7D0E484D414 CRC64;

  Query Match             34.1%;  Score 47;  DB 1;  Length 520;
  Best Local Similarity   39.3%;  Pred. No. 24;
  Matches   11;  Conservative    2;  Mismatches    5;  Indels   10;  Gaps    2;

Qy       1 RNPDG-DVGG---------PWAYTTNPR 18
           |:| | |: |           | || |||
Db     237 RDPGGSDIAGTTSTVQEQLEWIYTANPR 264


RESULT 41
TRD5_ECOLI
ID   TRD5_ECOLI     STANDARD;      PRT;   129 AA.
AC   P27192;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   TraD protein.
GN   TRAD.
OS   Escherichia coli.
OG   Plasmid IncP-beta R751.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-4.
RC   STRAIN=HB101;
RX   MEDLINE=92297959; PubMed=1818755;
RA   Miele L., Strack B., Kruft V., Lanka E.;
RT   "Gene organization and nucleotide sequence of the primase region of
RT   IncP plasmids RP4 and R751.";
RL   DNA Seq. 2:145-162(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Thomas C.M.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO PLASMID INCP-BETA RP4 TRAD.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59794; CAA42459.1; -.
DR   EMBL; U67194; AAC64471.1; -.
DR   PIR; S37670; S37670.
KW   Plasmid.
SQ   SEQUENCE   129 AA;  13600 MW;  A2F7296E4463BF88 CRC64;

  Query Match             33.7%;  Score 46.5;  DB 1;  Length 129;
  Best Local Similarity   50.0%;  Pred. No. 6.7;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       3 PDGDVGGPWAYTTNPRKLYD 22
           |  |:  |:  ||| |:|||
Db      93 PSADI-PPFITTTNARELYD 111


RESULT 42
POR2_HUMAN
ID   POR2_HUMAN     STANDARD;      PRT;   347 AA.
AC   P45880; Q9Y5I6;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Voltage-dependent anion-selective channel protein 2 (VDAC-2) (hVDAC2)
DE   (Outer mitochondrial membrane protein porin 2).
GN   VDAC2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-cell;
RX   MEDLINE=93280191; PubMed=7685033;
RA   Ha H., Hajek P., Bedwell D.M., Burrows P.D.;
RT   "A mitochondrial porin cDNA predicts the existence of multiple human
RT   porins.";
RL   J. Biol. Chem. 268:12143-12149(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93131931; PubMed=8420959;
RA   Blachly-Dyson E., Zambronicz E.B., Yu W.H., Adams V., McCabe E.R.,
RA   Adelman J., Colombini M., Forte M.;
RT   "Cloning and functional expression in yeast of two human isoforms of
RT   the outer mitochondrial membrane channel, the voltage-dependent anion
RT   channel.";
RL   J. Biol. Chem. 268:1835-1841(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 4).
RX   MEDLINE=99431679; PubMed=10501981;
RA   Decker W.K., Bowles K.R., Schatte E.C., Towbin J.A., Craigen W.J.;
RT   "Revised fine mapping of the human voltage-dependent anion channel
RT   loci by radiation hybrid analysis.";
RL   Mamm. Genome 10:1041-1042(1999).
RN   [4]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=95294003; PubMed=7539795;
RA   Yu W.H., Wolfgang W., Forte M.;
RT   "Subcellular localization of human voltage-dependent anion channel
RT   isoforms.";
RL   J. Biol. Chem. 270:13998-14006(1995).
CC   -!- FUNCTION: FORMS A CHANNEL THROUGH THE MITOCHONDRIAL OUTER
CC       MEMBRANE THAT ALLOWS DIFFUSION OF SMALL HYDROPHILIC MOLECULES.
CC       THE CHANNEL ADOPTS AN OPEN CONFORMATION AT LOW OR ZERO MEMBRANE
CC       POTENTIAL AND A CLOSED CONFORMATION AT POTENTIALS ABOVE 30-40 MV.
CC       THE OPEN STATE HAS A WEAK ANION SELECTIVITY WHEREAS THE CLOSED
CC       STATE IS CATION-SELECTIVE.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
CC       produced by alternative splicing and possibly alternative
CC       initiation.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED.
CC   -!- DOMAIN: CONSISTS MAINLY OF MEMBRANE-SPANNING SIDED BETA-SHEETS.
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC MITOCHONDRIAL PORIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L08666; AAA60144.1; -.
DR   EMBL; L08666; AAA60145.1; -.
DR   EMBL; L06328; AAB59457.1; -.
DR   EMBL; AF152227; AAD40241.1; -.
DR   EMBL; AF152220; AAD40241.1; JOINED.
DR   EMBL; AF152221; AAD40241.1; JOINED.
DR   EMBL; AF152222; AAD40241.1; JOINED.
DR   EMBL; AF152223; AAD40241.1; JOINED.
DR   EMBL; AF152224; AAD40241.1; JOINED.
DR   EMBL; AF152225; AAD40241.1; JOINED.
DR   EMBL; AF152226; AAD40241.1; JOINED.
DR   MIM; 193245; -.
DR   InterPro; IPR001925; Euk_porin.
DR   Pfam; PF01459; Euk_porin; 1.
DR   PRINTS; PR00185; EUKARYTPORIN.
```

```
DR   PROSITE; PS00558; EUKARYOTIC_PORIN; 1.
KW   Outer membrane; Porin; Mitochondrion; Alternative splicing;
KW   Alternative initiation; Polymorphism.
FT   VARSPLIC      1     26       MISSING (IN ISOFORM 2).
FT   VARSPLIC      1     26       MSWCNELRLPALKQHSIGRGLESHIT -> MATHGQTCAR
FT                                P (IN ISOFORM 3).
FT   VARSPLIC    303    347       SPWSWRLNPAERNLWEWISEDLALIYFHCDQQQAFFPPED
FT                                DQNKG -> LALEEA (IN ISOFORM 4).
FT   VARIANT      39     39       A -> V.
FT                                /FTId=VAR_006380.
FT   VARIANT      90     90       W -> C.
FT                                /FTId=VAR_006381.
SQ   SEQUENCE   347 AA;  38092 MW;  23F8ACA61184A1BD CRC64;

  Query Match             33.3%;  Score 46;  DB 1;  Length 347;
  Best Local Similarity   42.1%;  Pred. No. 22;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    2;  Gaps    1;

Qy       7 VGGPWAYTTNP--RKLYDY 23
           || ||::  ||  | |:::
Db     301 VGSPWSWRLNPAERNLWEW 319


RESULT 43
GAG_SIVAT
ID   GAG_SIVAT      STANDARD;      PRT;   519 AA.
AC   P05892;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   GAG polyprotein [Contains: Core proteins P17, P24, and P15].
GN   GAG.
OS   Simian immunodeficiency virus (TYO-1 isolate) (SIV-AGM).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11731;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88232906; PubMed=3374586;
RA   Fukasawa M., Miura T., Hasegawa A., Morikawa S., Tsujimoto H.,
RA   Miki K., Kitamura T., Hayami M.;
RT   "Sequence of simian immunodeficiency virus from African green monkey,
RT   a new member of the HIV/SIV group.";
RL   Nature 333:457-461(1988).
CC   -!- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.
CC   -!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X07805; CAA30657.1; -.
DR   PIR; A30045; FOLJG4.
DR   HSSP; P05888; 1AAF.
DR   HIV; X07805; GAG$AGMTY.
DR   InterPro; IPR000721; Gag_p24.
DR   InterPro; IPR000071; Retroviral_gag_p17.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00540; gag_p17; 1.
DR   Pfam; PF00607; gag_p24; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00234; HIV1MATRIX.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   AIDS; Core protein; Polyprotein; Zinc-finger; Repeat.
FT   CHAIN         1    141       CORE PROTEIN P17.
FT   CHAIN       142    519       CORE PROTEINS P24 AND P15.
FT   ZN_FING     397    414       CCHC-TYPE 1.
FT   ZN_FING     418    435       CCHC-TYPE 2.
SQ   SEQUENCE   519 AA;  58143 MW;  85A3AC06BCCDCA38 CRC64;

  Query Match             33.3%;  Score 46;  DB 1;  Length 519;
  Best Local Similarity   39.3%;  Pred. No. 34;
  Matches   11;  Conservative    2;  Mismatches    5;  Indels   10;  Gaps    2;

Qy       1 RNPDG-DVGG---------PWAYTTNPR 18
             |:| | |: |         | || |||
Db     237 RDPRGSDIAGTTSSVQEQLEWIYTANPR 264


RESULT 44
GAG_SIVAG
ID   GAG_SIVAG      STANDARD;      PRT;   521 AA.
AC   P27978;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   GAG polyprotein [Contains: Core proteins P17, P24, and P15].
GN   GAG.
OS   Simian immunodeficiency virus (AGM3 isolate) (SIV-AGM).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11730;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90232731; PubMed=2158689;
RA   Baier M., Garber C., Mueller C., Cichutek K., Kurth R.;
RT   "Complete nucleotide sequence of a simian immunodeficiency virus from
RT   African green monkeys: a novel type of intragroup divergence.";
RL   Virology 176:216-221(1990).
CC   -!- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.
CC   -!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M30931; AAA91913.1; -.
DR   HSSP; P03351; 1EIA.
DR   InterPro; IPR000721; Gag_p24.
DR   InterPro; IPR000071; Retroviral_gag_p17.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00540; gag_p17; 1.
DR   Pfam; PF00607; gag_p24; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00234; HIV1MATRIX.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   AIDS; Core protein; Polyprotein; Zinc-finger; Repeat.
FT   CHAIN         1    145       CORE PROTEIN P17.
FT   CHAIN       146    521       CORE PROTEINS P24 AND P15.
FT   ZN_FING     402    419       CCHC-TYPE 1.
FT   ZN_FING     423    440       CCHC-TYPE 2.
SQ   SEQUENCE   521 AA;  58409 MW;  1F111BD2F2EDF4F5 CRC64;

  Query Match             33.3%;  Score 46;  DB 1;  Length 521;
  Best Local Similarity   39.3%;  Pred. No. 34;
  Matches   11;  Conservative    2;  Mismatches    5;  Indels   10;  Gaps    2;

Qy       1 RNPDG-DVGG---------PWAYTTNPR 18
             |:| | |: |         | || |||
Db     241 RDPRGSDIAGTTSTVQEQLEWIYTANPR 268


RESULT 45
FA12_CAVPO
ID   FA12_CAVPO     STANDARD;      PRT;   603 AA.
AC   Q04962;
```

```
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE   (HAF) (Fragment).
GN   F12.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC   TISSUE=Liver;
RX   MEDLINE=93003367; PubMed=1390917;
RA   Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA   Kambara T., Okabe H.;
RT   "Primary structure of guinea-pig Hageman factor: sequence around the
RT   cleavage site differs from the human molecule.";
RL   Biochim. Biophys. Acta 1159:113-121(1992).
CC   -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC       THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC       GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC   -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC       VII to form factor VIIa and factor XI to form factor XIa.
CC   -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC       COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC       FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC       TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC       XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X68615; CAA48600.1; -.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.211; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Trypsin.
DR   InterPro; IPR000083; fibronectin_type_1.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW   Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     18
FT   CHAIN         19    358       ALPHA-FACTOR XIIA HEAVY CHAIN.
FT   CHAIN        359    603       ALPHA-FACTOR XIIA LIGHT CHAIN.
FT   DOMAIN        46     87       FIBRONECTIN TYPE-II.
FT   DOMAIN        93    130       EGF-LIKE 1.
FT   DOMAIN       132    172       FIBRONECTIN TYPE-I.
FT   DOMAIN       173    209       EGF-LIKE 2.
FT   DOMAIN       216    294       KRINGLE.
FT   DOMAIN       312    342       PRO-RICH.
FT   DOMAIN       359    603       SERINE PROTEASE.
FT   ACT_SITE     398    398       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     447    447       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     551    551       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID      97    109       BY SIMILARITY.
FT   DISULFID     103    118       BY SIMILARITY.
FT   DISULFID     120    129       BY SIMILARITY.
FT   DISULFID     134    162       BY SIMILARITY.
FT   DISULFID     160    169       BY SIMILARITY.
FT   DISULFID     177    188       BY SIMILARITY.
FT   DISULFID     182    197       BY SIMILARITY.
FT   DISULFID     199    208       BY SIMILARITY.
FT   DISULFID     216    294       BY SIMILARITY.
FT   DISULFID     237    276       BY SIMILARITY.
FT   DISULFID     265    289       BY SIMILARITY.
FT   DISULFID     345    472       BY SIMILARITY.
FT   DISULFID     383    399       BY SIMILARITY.
FT   DISULFID     391    461       BY SIMILARITY.
FT   DISULFID     422    425       BY SIMILARITY.
FT   DISULFID     488    557       BY SIMILARITY.
FT   DISULFID     520    536       BY SIMILARITY.
FT   DISULFID     547    578       BY SIMILARITY.
FT   CARBOHYD     248    248       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     270    270       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     419    419       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   603 AA;  66795 MW;  48DC6B946FB9ED59 CRC64;

  Query Match             33.3%;  Score 46;  DB 1;  Length 603;
  Best Local Similarity   38.1%;  Pred. No. 39;
  Matches    8;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       3 PDGDVGGPWAYTTNPRKLYDY 23
           |: |    |||   |:|: |:
Db      76 PNFDQDQQWAYCLEPKKVKDH 96


RESULT 46
YHX6_YEAST
ID   YHX6_YEAST     STANDARD;      PRT;   373 AA.
AC   P38866;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 42.4 kDa protein in CTR2-STB5 intergenic region.
GN   YHR176W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE=94378003; PubMed=8091229;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA   Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA   Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA   Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA   Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA   Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
RA   Vaudin M.;
```

```
RT   "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   VIII.";
RL   Science 265:2077-2082(1994).
CC   -!- SIMILARITY: BELONGS TO THE FMO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00027; AAB68021.1; -.
DR   PIR; S48915; S48915.
DR   SGD; S0001219; YHR176W.
DR   InterPro; IPR000960; FMO.
DR   Pfam; PF00743; FMO-like; 1.
DR   PRINTS; PR00370; FMOXYGENASE.
KW   Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
SQ   SEQUENCE   373 AA;  42440 MW;  DFB5BA11956896CB CRC64;

  Query Match             33.0%;  Score 45.5;  DB 1;  Length 373;
  Best Local Similarity   39.1%;  Pred. No. 28;
  Matches    9;  Conservative    5;  Mismatches    6;  Indels    3;  Gaps    1;

Qy       4 DGDVGGPWAY---TTNPRKLYDY 23
           | |:|| | |    :: | :||:
Db      40 DYDIGGVWHYPEQKSDGRVMYDH 62


RESULT 47
OASL_HUMAN
ID   OASL_HUMAN     STANDARD;      PRT;   514 AA.
AC   Q15646; O75686; Q9Y6K7; Q9Y6K6;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   59 kDa 2'-5'-oligoadenylate synthetase like protein (p59 OASL)
DE   (p59OASL) (Thyroid receptor interacting protein 14) (TRIP14).
GN   OASL OR TRIP14.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM P56).
RX   MEDLINE=98391734; PubMed=9722630;
RA   Hartmann R., Olsen H.S., Widder S., Joergensen R., Justesen J.;
RT   "p59OASL, a 2'-5' oligoadenylate synthetase like protein: a novel
RT   human gene related to the 2'-5' oligoadenylate synthetase family.";
RL   Nucleic Acids Res. 26:4121-4127(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS P56 AND P30).
RC   TISSUE=Monocytes;
RX   MEDLINE=99041549; PubMed=9826176;
RA   Rebouillat D., Marie I., Hovanessian A.G.;
RT   "Molecular cloning and characterization of two related and interferon-
RT   induced 56-kDa and 30-kDa proteins highly similar to 2'-5'
RT   oligoadenylate synthetase.";
RL   Eur. J. Biochem. 257:319-330(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   White S.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 260-416 FROM N.A. (ISOFORM P56).
RX   MEDLINE=95295737; PubMed=7776974;
RA   Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT   "Two classes of proteins dependent on either the presence or absence
RT   of thyroid hormone for interaction with the thyroid hormone
RT   receptor.";
RL   Mol. Endocrinol. 9:243-254(1995).
CC   -!- FUNCTION: DOES NOT HAVE 2'-5'-OAS ACTIVITY, BUT BINDS DOUBLE-
CC       STRANDED RNA AND DNA.
CC   -!- SUBUNIT: SPECIFICALLY INTERACTS WITH THE LIGAND BINDING DOMAIN OF
CC       THE THYROID RECEPTOR (TR). TRIP14 DOES NOT REQUIRE THE PRESENCE OF
CC       THYROID HORMONE FOR ITS INTERACTION.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; P56 (SHOWN HERE) AND P30; MAY BE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES, WITH THE HIGHEST
CC       LEVELS IN PRIMARY BLOOD LEUKOCYTES AND OTHER HEMATOPOIETIC SYSTEM
CC       TISSUES, COLON, STOMACH AND TO SOME EXTENT IN TESTIS.
CC   -!- INDUCTION: BY INTERFERONS.
CC   -!- SIMILARITY: BELONGS TO THE 2-5A SYNTHETASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 UBIQUITIN-LIKE DOMAINS.
CC   -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 386
CC       TO 416 DUE TO A FRAMESHIFT.
CC   -!- CAUTION: THIS MAY NOT BE THE TRUE ORTHOLOG OF MOUSE OASL.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ225089; CAA12396.1; -.
DR   EMBL; AF063611; AAD28541.1; -.
DR   EMBL; AF063612; AAD28542.1; -.
DR   EMBL; Z93097; -; NOT_ANNOTATED_CDS.
DR   EMBL; L40387; AAC41733.1; ALT_FRAME.
DR   HSSP; P02248; 1UBI.
DR   MIM; 603281; -.
DR   InterPro; IPR001797; 25A_synth.
DR   InterPro; IPR001201; PAP_25A_core.
DR   InterPro; IPR000626; Ubiquitin.
DR   Pfam; PF00240; ubiquitin; 2.
DR   SMART; SM00213; UBQ; 2.
DR   PROSITE; PS00832; 25A_SYNTH_1; 1.
DR   PROSITE; PS00833; 25A_SYNTH_2; 1.
DR   PROSITE; PS50152; 25A_SYNTH_3; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
KW   RNA-binding; Interferon induction; Alternative splicing; Repeat.
FT   DOMAIN      354    433       UBIQUITIN-LIKE 1.
FT   DOMAIN      434    509       UBIQUITIN-LIKE 2.
FT   VARSPLIC    220    255       YVKARSPRANLPPLYALELLTIYAWEMGTEEDENFM -> A
FT                                HHPGSGRPHPQRGRRVQMGHRCSEGLPVPETGLLL (IN
FT                                ISOFORM P30).
FT   VARSPLIC    256    514       MISSING (IN ISOFORM P30).
FT   CONFLICT     26     38       HREWKEEVLDAVR -> TGVEGRGARRCA (IN REF.
FT                                2).
FT   CONFLICT     89     89       S -> T (IN REF. 2; AAD28542).
FT   CONFLICT     95    113       QEAAKHHKDVLRLIWKTMW -> PGGSQASQRCSEADMENH
FT                                V (IN REF. 2).
FT   CONFLICT    223    223       A -> S (IN REF. 2).
FT   CONFLICT    317    317       Y -> I (IN REF. 2).
FT   CONFLICT    321    321       I -> T (IN REF. 2).
FT   CONFLICT    324    324       Q -> L (IN REF. 2).
FT   CONFLICT    341    342       NP -> KG (IN REF. 4).
FT   CONFLICT    445    445       S -> T (IN REF. 2; AAD28541).
SQ   SEQUENCE   514 AA;  59226 MW;  4D8BB655D9EA003E CRC64;

  Query Match             33.0%;  Score 45.5;  DB 1;  Length 514;
  Best Local Similarity   52.9%;  Pred. No. 39;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNP 17
           :|||   || :||  ||
Db     439 KNPD---GGSYAYAINP 452


RESULT 48
PPO_VITVI
ID   PPO_VITVI      STANDARD;      PRT;   607 AA.
```

```
AC   P43311;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Polyphenol oxidase, chloroplast precursor (EC 1.10.3.1) (PPO)
DE   (Catechol oxidase).
OS   Vitis vinifera (Grape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC   Vitis.
OX   NCBI_TaxID=29760;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SULTANA; TISSUE=Fruit;
RX   MEDLINE=95036022; PubMed=7948897;
RA   Dry I.B., Robinson S.P.;
RT   "Molecular cloning and characterisation of grape berry polyphenol
RT   oxidase.";
RL   Plant Mol. Biol. 26:495-502(1994).
CC   -!- FUNCTION: CATALYZE THE OXIDATION OF MONO- AND O-DIPHENOLS TO O-
CC       DIQUINONES.
CC   -!- CATALYTIC ACTIVITY: 2 catechol + O(2) = 2 1,2-benzoquinone + 2
CC       H(2)O.
CC   -!- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
CC   -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z27411; CAA81798.1; -.
DR   EMBL; A27657; CAA01887.1; -.
DR   InterPro; IPR002227; Tyrosinase.
DR   Pfam; PF00264; tyrosinase; 1.
DR   PRINTS; PR00092; TYROSINASE.
DR   PROSITE; PS00497; TYROSINASE_1; 1.
DR   PROSITE; PS00498; TYROSINASE_2; 1.
KW   Oxidoreductase; Copper; Metal-binding; Chloroplast; Transit peptide.
FT   TRANSIT       1    103       CHLOROPLAST (POTENTIAL).
FT   CHAIN       104    607       POLYPHENOL OXIDASE.
FT   METAL       211    211       COPPER A (BY SIMILARITY).
FT   METAL       220    220       COPPER A (BY SIMILARITY).
FT   METAL       342    342       COPPER B (BY SIMILARITY).
FT   METAL       346    346       COPPER B (BY SIMILARITY).
FT   METAL       375    375       COPPER B (BY SIMILARITY).
SQ   SEQUENCE   607 AA;  67347 MW;  B9045598E69BC57B CRC64;

  Query Match             33.0%;  Score 45.5;  DB 1;  Length 607;
  Best Local Similarity   47.6%;  Pred. No. 47;
  Matches   10;  Conservative    3;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       2 NPDGDVGGPWAYTTNPRKLYD 22
           |||| :  |  | ::|  |||
Db     249 NPDG-MYMPTIYASSPSSLYD 268


RESULT 49
NETR_MOUSE
ID   NETR_MOUSE     STANDARD;      PRT;   761 AA.
AC   O08762;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurotrypsin precursor (EC 3.4.21.-) (Motopsin) (Brain-specific serine
DE   protease 3) (BSSP-3).
GN   PRSS12 OR BSSP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97401523; PubMed=9245503;
RA   Gschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.;
RT   "Neurotrypsin, a novel multidomain serine protease expressed in the
RT   nervous system.";
RL   Mol. Cell. Neurosci. 9:207-219(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98008848; PubMed=9344839;
RA   Yamamura Y., Yamashiro K., Tsuruoka N., Nakazato H., Tsujimura A.,
RA   Yamaguchi N.;
RT   "Molecular cloning of a novel brain-specific serine protease with a
RT   kringle-like structure and three scavenger receptor cysteine-rich
RT   motifs.";
RL   Biochem. Biophys. Res. Commun. 239:386-392(1997).
CC   -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
CC       ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
CC       LEARNING AND MEMORY OPERATIONS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: MOST ABUNDANT IN CEREBRAL CORTEX, HIPPOCAMPUS
CC       AND AMYGDALA.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y13192; CAA73646.1; -.
DR   EMBL; D89871; BAA23986.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.237; -.
DR   MGD; MGI:1100881; Prss12.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001190; SRCR.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00530; SRCR; 3.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00202; SR; 3.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS00420; SRCR_1; 3.
DR   PROSITE; PS50287; SRCR_2; 3.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    761       NEUROTRYPSIN.
FT   DOMAIN       85    157       KRINGLE.
FT   DOMAIN      166    267       SRCR 1.
FT   DOMAIN      273    373       SRCR 2.
FT   DOMAIN      386    487       SRCR 3.
FT   DOMAIN      505    761       SERINE PROTEASE.
FT   DOMAIN      505    516       ZYMOGEN ACTIVATION REGION.
FT   ACT_SITE    516    517       REACTIVE BOND (POTENTIAL).
```

```
FT   ACT_SITE    562    562       CHARGE RELAY SYSTEM.
FT   ACT_SITE    612    612       CHARGE RELAY SYSTEM.
FT   ACT_SITE    711    711       CHARGE RELAY SYSTEM.
FT   DISULFID    505    636       POTENTIAL.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    521    521       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    569    569       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   761 AA;  84118 MW;  DF507B03712164E6 CRC64;

  Query Match             33.0%;  Score 45.5;  DB 1;  Length 761;
  Best Local Similarity   39.1%;  Pred. No. 59;
  Matches    9;  Conservative    4;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       1 RNPDGDVGGPWAYTTNPRKLYDY 23
           |:|||  | || :  | :   |:
Db     131 RSPDGS-GRPWCFYRNAQGKVDW 152


RESULT 50
FOLD_YEAST
ID   FOLD_YEAST     STANDARD;      PRT;   427 AA.
AC   Q12676;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Folylpolyglutamate synthase (EC 6.3.2.17) (Folylpoly-gamma-glutamate
DE   synthetase) (FPGS).
GN   FOL3 OR YMR113W OR YM9718.12.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=20261521; PubMed=10799479;
RA   Cherest H., Thomas D., Surdin-Kerjan Y.;
RT   "Polyglutamylation of folate coenzymes is necessary for methionine
RT   biosynthesis and maintenance of intact mitochondrial genome in
RT   Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 275:14056-14063(2000).
CC   -!- FUNCTION: CONVERSION OF FOLATES TO POLYGLUTAMATE DERIVATIVES.
CC   -!- CATALYTIC ACTIVITY: ATP + {tetrahydrofolyl-[Glu]}(N) + L-glutamate
CC       = ADP + phosphate + {tetrahydrofolyl-[Glu]}(N+1).
CC   -!- SIMILARITY: BELONGS TO THE FOLYLPOLYGLUTAMATE SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49702; CAA89750.1; -.
DR   HSSP; P15925; 1FGS.
DR   SGD; S0004719; FOL3.
DR   InterPro; IPR001645; FOLYLPOLYGLU_SYNT.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   PROSITE; PS01011; FOLYLPOLYGLU_SYNT_1; 1.
DR   PROSITE; PS01012; FOLYLPOLYGLU_SYNT_2; 1.
KW   Ligase; One-carbon metabolism; ATP-binding.
FT   NP_BIND     30     36       ATP (POTENTIAL).
SQ   SEQUENCE   427 AA;  47851 MW;  C3307CEF43BE1F30 CRC64;

  Query Match             32.6%;  Score 45;  DB 1;  Length 427;
  Best Local Similarity   37.5%;  Pred. No. 38;
  Matches    6;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       8 GGPWAYTTNPRKLYDY 23
           | ||  : |:| ::  |:
Db     356 GMPWIHATDPEEIKDF 371


Search completed: November  8, 2002, 09:31:41
Job time : 12 secs
```